HHS Public Access

Author manuscript

Ticks Tick Borne Dis. Author manuscript; available in PMC 2018 October 01.

Published in final edited form as:

Ticks Tick Borne Dis. 2017 October ; 8(6): 827–836. doi:10.1016/j.ttbdis.2017.06.008.

Transcriptional profiling of *Rickettsia prowazekii* coding and non-coding transcripts during *in vitro* host-pathogen and vector-pathogen interactions

Casey L. C. Schroeder[a,*], **Hema P. Narra**[a,*], **Abha Sahni**[a], **Kamil Khanipov**[b], **Jignesh Patel**[a], **Yuriy Fofanov**[b], and **Sanjeev K. Sahni**[a]

[a]Department of Pathology, University of Texas Medical Branch, 301 University Boulevard Galveston, TX 77555, USA

[b]Department of Pharmacology, University of Texas Medical Branch, 301 University Boulevard Galveston, TX 77555, USA

Abstract

Natural pathogen transmission of *Rickettsia prowazekii*, the etiologic agent of epidemic typhus, to humans is associated with arthropods, including human body lice, ticks, and ectoparasites of eastern flying squirrel. Recently, we have documented the presence of small RNAs in *Rickettsia* species and expression of *R. prowazekii* sRNAs during infection of cultured human microvascular endothelial cells (HMECs), which represent the primary target cells during human infections. Bacterial noncoding transcripts are now well established as critical post-transcriptional regulators of virulence and adaptation mechanisms in varying host environments. Despite their importance, little is known about the expression profile and regulatory activities of *R. prowazekii* sRNAs (*Rp*_sRs) in different host cells encountered as part of the natural life-cycle. To investigate the sRNA expression profile of *R. prowazekii* during infection of arthropod host cells, we employed an approach combining *in vitro* infection, bioinformatics, and PCR-based quantitation. Global analysis of *R. prowazekii* transcriptome by strand-specific RNA sequencing enabled us to identify 67 *cis*-acting (antisense) and 26 *trans*-acting (intergenic) *Rp*_sRs expressed during the infection of *Amblyomma americanum* (AAE2) cells. Comparative evaluation of expression during *R. prowazekii* infection of HMECs and AAE2 cells by quantitative RT-PCR demonstrated significantly higher expression of four selected *Rp*_sRs in tick AAE2 cells. Examination of the coding transcriptome revealed differential up-regulation of >150 rickettsial genes in either HMECs or AAE2 cells and yielded evidence for host cell-dependent utilization of alternative transcription start sites by 18 rickettsial genes. Our results thus suggest noticeable differences in the expression of both *Rp*_sRs as well as the coding transcriptome and the exploitation of multiple transcription

[*]indicates equal contribution to this work.

Conflict of interests

The authors declare no conflict of interest.

initiation sites for select genes during the infection of human endothelium and tick vector cells as the host and yield new insights into rickettsial virulence and transmission mechanisms.

Keywords

Rickettsia prowazekii; Small RNAs; RNA Sequencing; Vascular Endothelium; and Epidemic typhus

1. Introduction

Pathogenic bacteria in the Genus *Rickettsia* belong to two major groups, namely spotted fever and typhus, which continue to pose significant health threats to humans across the globe. *Rickettsia prowazekii*, the causative agent of epidemic typhus, is an obligate intracellular, Gram-negative bacterium transmitted primarily by the body lice (*Pediculus humanus corporis*). Consequently, outbreaks of epidemic typhus tend to occur with conditions of crowding in close quarters and compromised hygiene during the times of war, famine, or natural disasters and the disease is also known as camp/famine/jail fever. During human infections, vascular endothelial cells lining the small and medium-sized blood vessels are the primary targets of infection and salient features of disease pathogenesis include vascular inflammation/dysfunction and perturbation of the vasculature's barrier function manifesting as altered permeability and fluid leakage from the intravascular compartment to the interstitium (Bechah et al., 2008b; Walker and Ismail, 2008). Considered to be one of the most severe forms of human rickettsioses, epidemic typhus due to *R. prowazekii* is associated with high mortality rates, in particular during the absence of appropriate sanitation and timely intervention with antibiotics-based therapies (Bechah et al., 2008a; Raoult et al., 2004; Uchiyama, 2012). Also, the recrudescent form of epidemic typhus or Brill-Zinsser disease can manifest in patients years after the primary infection and clinical recovery (Bechah et al., 2008a; Uchiyama, 2012), and such a reoccurence can lead to new cases or outbreaks of epidemic typhus. Although human body lice are the established principal vectors, ectoparasites (fleas and lice) of the flying squirrel maintain *R. prowazekii* in the sylvatic cycle. The presence of *R. prowazekii* in *Amblyomma* ticks from Mexico and *Ixodes* ticks in the Netherlands has also been demonstrated recently, suggesting the possibility of tick transmission in natural infections (Bozeman et al., 1975; Medina-Sanchez et al., 2005; Philip et al., 1966).

Once thought to be junk DNA, bacteria encode small regulatory RNAs (sRNAs) that act as critical post-transcriptional regulators of gene expression. These sRNAs typically range from 50 to 500 nucleotides and regulate a variety of processes such as environmental sensing, metabolism, stress responses, and virulence in pathogenic bacteria. The major families of sRNAs include true antisense RNAs originating from the 'opposite' complementary strand to the mRNA (*cis*-acting), sRNAs that act by limited complementarity base pairing with their targets (*trans*-acting), and sRNAs that exhibit binding interactions with proteins affecting their activity. *Trans*-acting sRNAs are encoded within the intergenic regions and act on target RNAs located elsewhere in the genome. In essence, such sRNAs are akin to eukaryotic microRNAs in their ability to modulate the activity and stability of multiple

mRNAs (Gottesman and Storz, 2011; Liu and Camilli, 2010). Unlike *cis*-acting sRNAs, *trans*-acting sRNAs display only partial nucleotide complementarity and generally require an RNA chaperone to facilitate interactions with their targets (Waters and Storz, 2009).

Using a bioinformatics-based approach, we recently predicted the presence of over 1,700 *trans*-acting sRNAs in the genomes of 16 different strains encompassing 13 rickettsial species (Schroeder et al., 2015). Using infection of cultured human microvascular endothelial cells (HMECs) with *R. prowazekii*, we further identified expression of 35 novel *trans*-acting and 23 *cis*-acting sRNAs through Next Generation Sequencing and confirmed expression of four novel *R. prowazekii* sRNAs (named *Rp*_sRs) along with the highly conserved and known bacterial sRNAs, namely α-tmRNA, RNaseP_bact_a, *ffs*, and 6S RNA (Schroeder et al., 2016). The present study was undertaken to conduct a compare and contrast analysis of the expression of *R. prowazekii* transcriptome during infection of human and tick host cells. Our results enable identification of additional *Rp*_sR candidates uniquely expressed during infection of tick cells and suggest differential expression of select *Rp*_sRs in tick vis-à-vis human host cells. In addition, a comprehensive analysis of encoding rickettsial transcriptome during infection of human and tick vector host cells yields the first evidence for the utilization of multiple transcription start sites depending on the host niche.

2. Materials and Methods

2.1. Rickettsia prowazekii and Cell Culture

Stocks of *R. prowazekii* strain Breinl were prepared by infecting Vero cells cultured in DMEM supplemented with 2% fetal bovine serum in an atmosphere of 95%O_2: 5% CO_2 at 35°C following standard protocols (Rydkina et al., 2005). Rickettsiae were purified by differential centrifugation, titered by a quantitative PCR-based assay, and stored at −80°C as aliquots until use (Labruna et al., 2004). Considering that repeat freeze-thaw cycles may alter rickettsial viability and transcriptome, all experiments were performed using *R. prowazekii* stocks gently thawed on ice for the first time. Human dermal microvascular endothelial cells (HMECs) were cultured at 37°C with 5% CO_2 in MCDB131 medium supplemented with 10% fetal bovine serum, 10mM L-glutamine, 1μg/ml hydrocortisone, and 10ng/ml epidermal growth factor as previously described (Schroeder et al., 2015; Schroeder et al., 2016). The use of human cell lines in our study was exempt by the University of Texas Medical Branch (UTMB) Institutional Review Board (IRB), but approved by the UTMB Institutional Biosafety Committee (IBC). *Amblyomma americanum* tick cells (AAE2) were grown in L-15B complete medium (pH 7.5) at 34°C to ~90% confluence. Approximately 24h prior to infection, the medium in each flask was replaced with L-15B infection medium (pH 7.5) containing 25mM sodium bicarbonate and HEPES (Munderloh and Kurtti, 1989). *In vitro* infection of HMECs with *R. prowazekii* stocks was carried out at 37°C or 34°C according to our standard protocols and procedures (Rydkina et al., 2005; Narra et. al., 2016). To achieve a comparable level of infection, AAE2 cells were infected with different stocks of *R. prowazekii* and incubated at 34°C. At 24 h, the cells were gently scraped and pelleted by centrifugation at 400 g for 5 minutes. The cell pellet was washed twice with sterile phosphate buffered saline (PBS) and processed for total DNA extraction using DNeasy Blood & Tissue kit (Qiagen). The MOI was estimated by absolute

quantification using gene specific primers (Supplementary table 1) targeting tick calreticulin and *R. prowazekii* citrate synthase (*gltA*) genes. For RNA-Seq experiments, HMECs were infected with *R. prowazekii* at an MOI of 5:1 (~6×10^4 pfu of rickettsiae per cm^2) in minimal volume of MCDB131 medium and incubated at room temperature for 15 minutes with gentle, intermittent rocking to enhance adhesion and invasion. The inoculation medium was then replaced with fresh medium and the cells were incubated for 24 h at 37°C, 5% CO_2 (Schroeder et al., 2015; Schroeder et al., 2016). For AAE2 cells, the L-15B medium (containing viable, semi-adherent cells) from the culture flask was collected to pellet the non-adherent cells by centrifugation at 400 g for 5 minutes. The pellet was then suspended in 1 ml of L-15B infection medium and added back to the adherent cells within the same flask. The cells were then infected with *R. prowazekii* at an MOI of 5:1 based on the estimation of infectivity titers as described above and gently rocked at room temperature for 15 minutes, at which time the medium containing rickettsiae was replaced with fresh medium as described above and the flasks were placed in an incubator at 34°C for 24 h. At the end of incubation, the medium was completely removed and total RNA was extracted by the Tri-Reagent method detailed below. Each RNA-Seq experiment was performed on two independent biological replicates.

2.2. RNA Isolation and Sequencing

Isolation of total RNA from HMECs and AAE2 cells infected with *R. prowazekii* was carried out at 24h using our standard Tri-Reagent (Molecular Research Center) protocol for deep sequencing. Total RNA was treated with DNaseI (Zymo Research) to remove any contaminating genomic DNA, and the samples were further processed using Dynabeads® Oligo $(dT)_{25}$ (ThermoFisher Scientific) and Ribo-Zero (Epicentre) to remove any interfering eukaryotic mRNAs and ribosomal RNAs, respectively. The enriched total RNA preparations thus obtained were quantified using the MultiSkan Go Microplate Spectrophotometer (ThermoScientific) and assessed for their quality on an Agilent 2100 Bioanalyzer (Agilent Technologies).

Subsequently, the total RNA from each biological replicate was divided into two equal aliquots. One aliquot was treated with Terminator 5′-Phosphate-Dependent Exonuclease (TEX) (Epicentre) resulting in the degradation of processed RNA transcripts containing the 5′ monophosphate (+TEX) but not the primary transcripts with 5′ triphosphate. The other aliquot served as untreated control and contained both processed and primary transcripts (−TEX). Independent cDNA libraries for each aliquot were generated using the TruSeq RNA Sample Prep Kit (Illumina) as per manufacturer's directions. Strand-specific sequencing on non-size selected cDNA libraries was performed on an Illumina HiSeq 1500 at our institutional Next Generation Sequencing Core facility. The sequencing libraries were comprised of 100 base long reads in a FASTQ format. The quality of each read was assessed and any base with a PHRED < 15 was excluded from the analysis. The first 14 bases of each read were trimmed and the remaining 86 base long reads were mapped onto the *R. prowazekii* Breinl genome (NC_020993) allowing up to four base mismatches using Bowtie2 (Langmead and Salzberg, 2012). Reads that did not map to the *R. prowazekii* Breinl genome or that mapped to more than one genomic location were discarded from the analysis. 'Transcripts per million' (TPM) is computed as:

$$\frac{RPKM * 10^6}{\sum RPKM}$$

where the sum is over the 'reads per kilobase per mapped reads' (RKPM) values of all genes/transcripts (Li et al., 2010). The RKPM is measured as shown below and described earlier (Mortazavi et al., 2008)

$$\frac{total\ reads\ mapping\ to\ the\ gene\ (ORF)}{mapped\ reads\ (millions) * gene\ length\ (KB)}$$

2.3. Quantitative Real-Time Reverse Transcriptase PCR (q-RT-PCR)

To validate the expression profile of differentially expressed *Rp*_sRs and *R. prowazekii* genes, q-RT-PCR was performed on total RNA extracted from HMECs and AAE2 cells infected with *R. prowazekii*. Because temperature shifts are known to alter the rickettsial transcriptome (Dreher-Lesnick et al., 2008; Ellison et al., 2009; Galletti et al., 2013), we performed q-RT-PCR on *R. prowazekii* infected HMECs at both 37°C and 34°C and on infected AAE2 cells grown at 34°C. HMECs and AAE2 cells were infected with *R. prowazekii* as described in Section 2.1 above. Briefly, after gentle rocking for 15 minutes at room temperature, the medium in the flasks was replaced with fresh medium and infected HMECs were incubated at either 37°C or 34°C with 5% CO_2 and AAE2 cells infected with *R. prowazekii* were incubated at 34°C. Quantitative RT-PCR was performed on RNA extracted at 3h and 24h post-infection to capture the expression profiles of *R. prowazekii* small RNAs and transcripts at early (3h) and late (24h) stages of infection. For comparative analysis, HMECs and AAE2 cells infected for 30 minutes at the respective temperatures were designated as the 'base line'. Total RNA was extracted by Tri-reagent method following our standard protocol. One microgram (1μg) of DNaseI treated total RNA was reverse transcribed using High Capacity Reverse Transcription Kit (Applied Biosystems) with random primers following the manufacturer's instructions. q-RT PCR was performed using SYBR® Green PCR Master Mix (Life Technologies). Each 20 μL reaction contained 1X SYBR® Green PCR Master Mix (contains DNA polymerase and dNTPs), 0.5 μM forward primer, 0.5 μM reverse primer, and 100 ng cDNA template. Thermal cycler conditions were: stage 1 at 95°C for 10 minutes, stage 2 (40 cycles) at 95°C for 30 seconds and 60°C for 30 seconds, followed by melt curve. *R. prowazekii* 16S rRNA was used as internal control and the expression of *Rp*_sR and rickettsial transcripts was analyzed by $2^{-\Delta\Delta Ct}$ method (Schmittgen and Livak, 2008). The data are presented as the mean±SEM from a minimum of three biological replicates processed as two technical replicates. All primers used in this study are listed in Supplementary table 1. Statistical analysis was performed using GraphPad Prism using Mann-Whitney t-test with statistical significance set to a threshold *P*-value of ≤ 0.05.

2.4. Computational Identification of Target Genes Regulated by Rp_sRs

To gain insight into the post-transcriptional regulation by *Rp*_sRs, targets genes potentially regulated by *Rp*_sR76, 83, 86 and 159 were predicted using IntaRNA algorithm (Busch et al., 2008) using default parameters with the only exception that the region of interrogation for identification of seed region was set to +150 and −100 bases with respective to the translational start codon of an ORF (Narra et. al., 2016). Only target genes exhibiting a significance of $p<0.05$ for the *Rp*_sR-target mRNA seed region interaction were considered for the analysis.

3. Results

3.1. R. prowazekii Coding Sequence (CDS) Expression in HMECs and AAE2 Cells

Global transcriptional profiling is a comprehensive approach to better understand the expression profile of coding and non-coding transcripts in a given condition. To gain functional insights into the transcriptional landscape of *R. prowazekii* during its interaction with host vis-à-vis vector cells *in vitro*, we compared the expression profiles of *R. prowazekii* coding genes (ORFs) in infected HMECs and *A. americanum* tick (AAE2) cells via normalization of our RNA-seq data by calculating the transcripts per kilobase million (TPM) for each rickettsial ORF (Supplementary table 2). This method was chosen over the "reads per kilobase per million" (RPKM), because TPM eliminates an intrinsic statistical bias attributed to an inconsistent measure of molar concentrations (Wagner et al., 2012). The comparative analysis of the expression levels of *R. prowazekii* ORFs in HMECs versus AAE2 cells is presented in Supplementary table 2. Employing a cut-off of 3-fold or higher, we determined that 34 rickettsial genes were expressed at significantly higher levels in HMECs (Table 1). Among these, a majority (82%) were found to be expressed at levels 3 to 7 times higher than in AAE2 cells. The most striking change of 11 to 23 times higher expression levels in relation to AAE2 cells was noticed for three genes namely, *H375_9040* (Heat shock protein 60), *H375_7520* (Tol-Pal system peptidoglycan lipoprotein), and *H375_6910* (hypothetical protein). Conversely, about 72% of rickettsial genes were found to be expressed in AAE2 cells at levels 3 to 7 times greater than in HMECs and there was very high expression of another 16 rickettsial genes in AAE2 cells as evidenced by an increase of 10 to 29-fold over the same during the infection of HMECs (Table 2). Intriguingly, the steady-state levels of transcripts for a total of 47 rickettsial genes were below the limit of detection in HMECs despite their transcription during infection of AAE2 cells, whereas transcripts for only four genes could not be detected in AAE2 cells despite their expression in HMECs. Furthermore, expression levels of a total of 11 genes were below the limit of detection in either host cell type (Supplementary table 2). Of these, 10 genes were annotated as the hypothetical proteins with as yet unassigned functions and the remaining gene, *H375_8360*, encodes for a lipid A export ATP-binding/permease protein.

3.2. q-RT-PCR based quantitation of Novel Differentially Expressed R. prowazekii genes in Human and Tick cells

To validate the differential expression of *R. prowazekii* transcripts observed in our RNA Seq, q-RT PCR was performed on four genes, namely, *H375_9040* and *H375_7520* (upregulated during infection of HMECs), and *H375_8480* and *H375_1050* (highly expressed during

infection of AAE2 cells). As temperature shifts are known to alter rickettsial gene expression, total RNA preparations from HMECs infected with *R. prowazekii* at either 37°C or 34°C and AAE2 cells infected at 34°C were used for comparative quantitative analysis using gene-specific primers and 16S rRNA as an internal control. Interestingly, *H375_9040* and *H375_7520* were highly expressed during the infection of HMECs, whereas *H375_8480* and *H375_1050* were up-regulated during tick cell infection (Figure 1). Quantitative RT-PCR based analysis of these genes exhibited an overall pattern of expression consistent with the RNA-Seq data, although some variation in fold-change values were noted as expected. Notably, all genes exhibited highest transcript abundance at 24 h post-infection in both cell types. Although temperature (37°C vs. 34°C) had an influence on the expression of *H375_1050* during the infection of HMECs, significant differences were also evident in the transcript abundance depending on the host niche, thus confirming the results from global transcriptomic profiling of *R. prowazekii* transcripts during host-pathogen and vector-pathogen interactions *in vitro*.

3.3. Alternate ORF Transcription Start Sites during the Infection of Human and Tick Cells

Regulation of transcription initiation during mRNA biogenesis represents the first layer in the control of gene expression and alternative transcription initiation results in the generation of transcripts differing in the length of the 5′-untranslated region (5′-UTR). Identification of mRNA transcription start sites (TSSs) is, therefore, critical for characterization of promoter regions, which is essential for understanding gene expression and regulation patterns in both prokaryotes as well as eukaryotes. Application of differential RNA-seq relying on the tri-phosphorylation of primary, but not processed, transcripts at their 5′ ends has led to the mapping of TSSs in *E. coli* grown under different conditions (Wade, 2015). To examine possible differences in rickettsial TSS during infection of human and tick cells as the host, we subjected our RNA samples to treatment with terminator 5′-phosphate-dependent exonuclease (TEX) prior to sequencing. Under these experimental conditions, an average of 333,431 and 256,084 reads belonging to primary transcripts mapped to *R. prowazekii* genome in HMECs and AAE2 cells, respectively, and allowed us to identify TSSs for 97 rickettsial genes expressed during the infection of host and tick cells (Supplementary table 3). Further in-depth examination of sequencing data from *R. prowazekii*-infected HMECs and AAE2 cells revealed utilization of an alternative TSS by 18 rickettsial genes involved in diverse set of rickettsial pathways, such as transcription, metabolism, signaling and type I secretion system (Table 3). A strand bias was also apparent in that 12 of 18 genes were located on the positive strand. Interestingly, the transcript length was longer for a majority of these genes (12 of 18) during their expression in AAE2 cells (Table 3).

3.4. Identification of R. prowazekii small RNAs expressed during the infection of AAE2 tick cells

We have recently demonstrated that *R. prowazekii* encodes and expresses a number of *Rp*_sR's during *in vitro* infection of human microvascular endothelial cells (Schroeder et al., 2015). To investigate the possibility of divergence in sRNA transcriptome of *R. prowazekii* in different host environments, we infected AAE2 cells for in-depth analysis by next generation sequencing to identify and catalogue *Rp*_sR's specifically expressed during vector-pathogen interactions. To achieve a direct comparison of *Rp*_sR profiles during the

infection of human and tick cells as the host, we carried out a side-by-side analysis of infected HMECs as well. RNA sequencing of cDNA libraries prepared from AAE2 cells and HMECs infected with *R. prowazekii* and enriched for bacterial transcripts resulted in approximately an average of 91 million and 46 million total reads, respectively. Despite enrichment of total RNA preparations for bacterial transcripts, a significant percentage of reads corresponded to the eukaryotic genome and mapped to non-polyadenylated transcripts originating from mitochondrial genes, rRNAs, tRNAs, and long and small non-coding RNAs. In correlation with our previous study (Schroeder et al., 2016), our sequencing analysis resulted in approximately an average of 2.9 million and 1.3 million reads mapping to the bacterial genome in libraries originating from total RNA isolated from AAE2 cells and HMECs infected with *R. prowazekii*, respectively. Overall, the eukaryotic and bacterial contributions in our RNA sequencing is in general agreement with the finding that obligately intracellular bacterial genomes only constitute 2–5% of extracted total RNA despite application of efficient enrichment protocols and that only 5% of the bacterial RNA represents mRNAs and sRNAs, with the remaining 95% mapping to rRNAs and tRNAs (Westermann et al., 2012). Thus, application of genome-wide identification of sRNA candidate-containing regions within the RNA-Seq datasets revealed an additional 93 novel candidate sRNA-encoding regions in intergenic regions (Figure 2, Supplementary table 4) and transcripts from regions antisense to open reading frames bearing the characteristics of *cis*-acting antisense sRNAs at 24h post-infection (Figure, Supplementary table 4). Of these, *cis*-acting accounted for approximately 72% of newly identified sRNAs in AAE2 cells, among which *Rp*_sR152, a *cis*-acting sRNA antisense to *H375_8370*, was the smallest with a length of 148 bp. Interestingly, 11 other sRNAs (*Rp*_sR79, *Rp*_sR92, *Rp*_sR93, *Rp*_sR94, *Rp*_sR114, *Rp*_sR120, *Rp*_sR128, *Rp*_sR130, *Rp*_sR135, *Rp*_sR143, and *Rp*_sR154) ranged from 516 bp to 844 bp in length and 8 of these were classified as *cis*-acting. No strand bias in regards to their origin was evident based on the expression of 47% of *Rp*_sRs from the leading strand and remaining 53% on the lagging strand. There were no significant differences in strand-specificity as well based on the location of 38% *trans*-acting and 50% *cis*-acting *Rp*_sRs on the leading strand. Further, 6 (*Rp*_sR71, *Rp*_sR74, *Rp*_sR75, *Rp*_sR129, *Rp*_sR139, and *Rp*_sR155) of the 26 *trans*-acting *Rp*_sRs were encoded in the same orientation as their upstream genes. As expected, 66 *cis*-acting sRNAs were found to be directly anti-sense to an open reading frame on the opposite strand. Interestingly, *Rp*_sR124 was expressed in a manner such that it overlapped with the 3′ end of a 3-polyprenyl-4-hydroxybenzoate carboxylase (*H375_7030*) and the 5′ end of *H375_7040* coding for a polyhydroxyalkanoic acid synthase. Collectively, our results reveal the expression of several novel *Rp*_sRs during *R. prowazekii* infection of tick cells.

3.5. q-RT-PCR based quantitation of Novel Differentially Expressed sRNAs in Human and Tick cells

As previously reported, bacterial sRNA expression can not only differ among tissues and organ systems within an infected host, but also between genders of a particular host species (Woolfit et al., 2015) and evidence from our laboratory has suggested differential regulation of *R. conorii* sRNAs in tick and human host cells (Narra et al., 2016). To further investigate whether *R. prowazekii* selectively and/or differentially express sRNAs in different host cells, we measured the expression levels of five novel candidate sRNAs in HMECs infected at

either 37°C or 34°C and AAE2 cells infected at 34°C. The steady state levels of *Rp*_sRs in cells infected for only 0.5h served as the baseline for comparative analysis at 3 and 24h post-infection using 16S rRNA as the endogenous control. Both *Rp*_sR76 and *Rp*_sR83 demonstrated significantly higher levels of expression in AAE2 cells at 3h post-infection when compared to HMECs grown at 37°C ($p \leq 0.001$), but failed to show any significance differences at the later time of 24h. Interestingly, significant difference ($p < 0.05$) in *Rp*_sR76 and *Rp*_sR83 expression was seen at 24h between HMECs and AAE2 cells grown at 34°C (Figure 4). Consistent with our RNA Seq data, *Rp*_sR86 was significantly highly expressed during AAE2 cell infection at both early (3h) and late (24h) time points when compared to HMECs grown at either 37°C or 34°C. *Rp*_sR159 also displayed significantly higher expression in AAE2 cells at 3h ($p \leq 0.01$) when compared to the respective expression levels in HMECs at both 37°C and 34°C. The highest level of *Rp*_sR159 expression was, however, noticed in HMECs at 34°C during late stage of infection (24h), suggesting the potential effects of temperature on its expression (Figure 4). Also, there was no evidence of any significant differences in the pattern of expression of *Rp*_sR74 during *R. prowazekii* infection of HMECs and AAE2 cells (data not shown).

3.6. Prediction of target genes regulated by Rp_sR's

Using IntaRNA, we have predicted a total of 45, 27, 38 and 36 genes as the potential targets of regulation by *Rp*_sR76, 83, 86 and 159, respectively. Of these, a number of important genes such as diguanylate cyclase (*H375_RS01895*), protease modulator *hflC* (*H375_RS02485*), ferredoxin-NADP(+) reductase (*H375_RS00510*), resolvase (*H375_RS03980*), and two hypothetical proteins (*H375_RS02580* and *H375_RS03250*) may be regulated by at least two sRNAs. In addition, several other key genes, including multidrug transporter (*H375_RS00115*), two component sensor histidine kinase (*H375_RS04485*), DNA repair (*H375_RS02190*), and cold shock protein *cspA* (*H375_RS04205*) emerged as the potential regulatory targets of at least one *Rp*_sR (Supplementary table 5).

4. Discussion

In-depth transcriptomic analysis offers a comprehensive mapping and enhanced understanding of the systems biology of host-pathogen interactions. In the present study, we have employed an RNA-sequencing based approach to perform comparative evaluation of the coding as well as non-coding transcriptome of *R. prowazekii* during the infection of human endothelial cells as the preferred, primary target cell type in the mammalian host and tick vector cells as the arthropod stage of the pathogen's life-cycle. Although body lice are the major established natural vectors responsible for epidemic typhus outbreaks, the rationale for our comparative analysis derives from recent reports documenting the presence of *R. prowazekii* in circulating tick vectors and easy availability as well as infectivity of cultured tick cells. Because RNA-seq studies using deep sequencing technologies have unequivocally enhanced our understanding of the extent and complexity of both prokaryotic and eukaryotic transcriptomes, we focused our investigation of rickettsial transcriptome on all transcript species, including mRNAs and non-coding RNAs, with the ultimate goal of determining the characteristics of genes in terms of their usage of transcriptional start sites and quantifying the changes in expression levels of each transcript during pathogen interplay

with human *versus* tick host cells (Sharma and Vogel, 2014). Based on the genomic location and transcript orientation in relation to that of adjoining upstream and downstream genes and the number of mapped reads, we report on the presence of 93 sRNA candidates abundantly expressed in AAE2 cells in addition to 70 small RNAs (*Rp*_sR's) previously identified during the infection of HMECs. Subsequent analysis of a few selected candidates, known to be selectively and differentially expressed in AAE2 cells and potentially regulating target genes involved in signaling, transport, cold shock and DNA repair (Supplementary table 5), by quantitative PCR further ascertains that *Rp*_sR76, *Rp*_sR83, *Rp*_sR86, and *Rp*_sR159 are expressed at significantly higher levels in arthropod cells, suggesting the possibility of an important role for these newly-identified candidate RNAs in pathogen interactions with and survival in arthropod cells. Because temperature shifts are known to alter the rickettsial transcriptome (Dreher-Lesnick et al., 2008; Ellison et al., 2009; Galletti et al., 2013), we assessed the expression of *Rp*_sRs in HMECs grown at either 37°C or 34°C and AAE2 cells grown at 34°C to delineate the impact of temperature *versus* host niche on sRNA expression. Interestingly, temperature variation had little effect on the transcript abundance of three *Rp*_sR's (*Rp*_sR76, 83, and 86), while only *Rp*_sR159 showed significant differences in its expression dependent on the temperature and host niche (Figure 4). These results are in congruence to recently emerging evidence implicating differential regulation of sRNAs in different strains of a vertically transmitted, endosymbiotic α-proteobacterium *Wolbachaia pipientis* and under specific environmental conditions such as the infection of different host tissues and sexes. Specifically, a conserved intergenic sRNA *ncrwmel02* is reportedly expressed at two and seven times higher levels, respectively, in strains wMel and wAu of *W. pipientis*, when compared to strains wMelCs and wMelPop. Furthermore, *ncrwmel02* is present at significantly higher levels in the abdomens of male *Drosophila melanogaster* as their natural hosts than in the abdomens of female flies and there is more than 10-fold higher expression in testes compared to ovaries, suggesting both tissue-specific and host sex-specific regulation (Woolfit et al., 2015). Similarly, a broad analysis of *Burkholderia thailandensis* sRNA expression profiles via microarrays covering intergenic regions of more than 90 bases suggests differential expression of 38 novel and 2 conserved small RNAs in response to varying stressors (Stubben et al., 2014) and a comprehensive transcriptomic analysis of *Burkholderia pseudomallei* exposed to diverse physical, chemical, and biological conditions also reveals context-dependent expression of non-coding sRNAs, including a number of *cis*-regulatory motifs (Ooi et al., 2013).

Although a number of bacterial regulatory RNAs have been classified as *trans*-encoded sRNAs which require the RNA binding chaperone protein Hfq to facilitate base pairing with their target mRNAs (Khandige et al., 2015), about half of all Gram-positive and Gram-negative bacteria, including *Rickettsia* species, do not encode for Hfq (Dugar et al., 2013; Östberg et al., 2004). Recently, we have shown *R. conorii* trans-acting sRNA *Rc*_sR42 to directly bind to *cydA* mRNA *in vitro*, implicating the possibility of a chaperone-independent mechanism of sRNA-mRNA interaction in *Rickettsia* species (Narra et al., 2016). Using bioinformatics based approach, this study further identified several target genes involved in signaling, DNA repair, and cold shock response to be regulated by *Rp*_sR's selectively and differentially expressed during tick cell infection. Further experimental investigations to

validate the computational target gene predictions of rickettsial sRNAs are necessary and currently in progress.

It is believed that Gram-positive bacteria, in general, rely more on *cis*-acting regulatory mechanisms such as riboswitches than Gram-negative organisms known to utilize *trans*-acting sRNAs much more extensively (Lasa et al., 2012; Waters and Storz, 2009). Recent evidence suggests that approximately 10% of the genes in the environmental bacterium *B. pseudomallei* are subject to regulatory control by antisense transcription (Ooi et al., 2013). Similarly, a global transcriptomic analysis of *Helicobacter pylori*, a Gram-negative, *ε*-proteobacterium, has revealed at least one antisense transcription start site for nearly 46% of all ORFs and 28% of tRNAs (Sharma et al., 2010), whereas antisense transcription for nearly 50% of the coding ORFs has also been reported for *Staphylococcus aureus* as a Gram-positive bacterial pathogen (Lasa et al., 2011). In this regard, an intriguing aspect of our findings is that about 25% of *R. prowazekii* sRNAs expressed in AAE2 cells are categorized as *trans*-acting. Thus, combining the repertoire of novel sRNA candidates in AAE2 cells with those previously identified during the infection of HMECs allows for the classification of approximately 44% as *trans*-acting and 55% as *cis*-acting, representing a ratio closer to other reported organisms (Supplementary table 4).

Cis-encoded sRNAs overlapping functionally defined genes are pervasive through the genomes of prokaryotes and such antisense RNAs have been proposed to play a role analogous to that of transcription factors in transcription regulation in adaptive transition between distinct states. In the present study, we have identified 67 *cis*-acting sRNAs antisense to key ORFs encoding for structural proteins, transporters, membrane lipoproteins, and pathways of metabolism. For example, *Rp*_sR101 and *Rp*_sR102 are antisense to a VirB6 paralog (*H375_5210*) and the ATPase VirB4 (*H375_5270*) of the *vir*-induced type IV secretion system (T4SS), respectively. Spanning multiple membranes, T4SS is a complex multi-protein transporter encoded in many Gram-negative bacteria, which forms a syringe-like apparatus that functions to deliver a variety of virulence factors into host cells (Gillespie et al., 2009; Gillespie et al., 2015). The T4SS is composed of at least 12 Vir proteins with multiple paralogs performing unique functions as part of the overall complex (Gillespie et al., 2016). The VirB6 component, which composes the inner channel of T4SS essential for substrate transfer, is among the most divergent of the VirB proteins with at least five paralogs (Gillespie et al., 2009). Likewise, the VirB4 protein is an integral part of the T4SS due to its ATPase activity, which provides the required energy for operation. Similarly, *cis*-acting *Rp*_sR148 is present antisense to the gene coding for Outer membrane protein B (H375_8270) and an ATP-dependent helicase encoding UvrD (*H375_1620*) may hypothetically be regulated by *Rp*_sR77 as an antisense sRNA. With regards to their function(s), as an abundantly expressed protein expressed on the surface of all *Rickettsia* species, Outer membrane protein B (also known as Sca5) is involved in rickettsial binding and invasion of eukaryotic host cell and antibodies directed against OmpB have been reported to protect mice from lethal doses of rickettsial infection (Chan et al., 2009). As a component of the nucleotide excision repair and the transcription coupled repair machinery, UvrD acts in concert with UvrC to excise dimerized nucleotides for final repair by DNA polymerase I (Van Houten and Kad, 2014).

Bacteria are known to respond to environmental cues through a network of regulatory RNAs and proteins as the determinants of genome-wide transcription patterns. Many of such regulatory mechanisms depend on the initiation of messenger RNA synthesis by RNA polymerase at the transcription start sites. Accordingly, location of TSS and quantitative determination of changes in TSS usage is an important step to understand bacterial gene regulation. Although our fundamental understanding of basic mechanisms of transcription activation has arisen from the investigations of simple promoters such as *lac* and *gal* in *E. coli*, a majority of activator-dependent promoters are much more complex due to co-regulation either by another activator or repressor or possibly by both (Barnard et al., 2004). Such naturally occurring promoters allow bacteria to respond rapidly to specific environmental conditions. As an example, *Salmonella* regulates flagellar transcription through multiple promoters based on specific environmental conditions (Wozniak et al., 2010) and global examination of transcription start sites in *Caulobacter crescentus,* an α-proteobacterium closely related to *Rickettsia*, reveals the origin of transcription of 53 of its 769 genes from multiple start sites (McGrath et al., 2007). In *Orientia* (formerly *Rickettsia) tsutsugamushi,* there is evidence for the use of tandem promoters for the production of 56-kDa type-specific antigen as an abundant surface protein and two independent promoter-like sequences upstream from transcription start points for citrate synthase (*gltA)* transcripts are encoded in *R. prowazekii* (Cai et al., 1995; Ohashi et al., 1992). Identified in this study, the genes utilizing multiple TSSs encode for proteins essential to rickettsial metabolism, secretion, and other housekeeping functions. Upon examination, four genes with a >3 TPM difference between AAE2 or HMECs exhibit multiple TSSs, of which *H375_2140*, *H375_3390*, and *H375_3940* show higher expression in AAE2 cells and *H375_4210* displays higher expression in HMECs. Interestingly, two genes namely *H375_2140* and *H375_3390* are annotated as a conserved rickettsial protein and hypothetical protein, respectively. Nearly 14% of *Methanolobus psychrophilus* R15 genes are shown to have an alternate TSS, resulting in the generation of unique transcript isoforms under cold responsive conditions (Li et al., 2015). We propose that alternate TSSs identified in this study may have an impact on the translational efficiency and mRNA stability of the coding gene. Further investigations into these genes should yield new clues pertaining to transcriptional regulation during pathogen interactions with the human and vector host cells as the supportive intracellular milieu for rickettsiae.

Microarray-based screens have identified several rickettsial genes to be differentially expressed upon a shift in temperature and during natural blood feeding (Dreher-Lesnick et al., 2008; Ellison et al., 2009; Galletti et al., 2013). In this study, we decipher the global transcriptional landscape of *R. prowazekii* during host-pathogen and vector-pathogen interactions *in vitro*. Follow up quantitative analysis of the expression of four rickettsial genes encoding for heat shock protein 60 family co-chaperone GroES (*H375_9040*), Tol-Pal system peptidoglycan-associated lipoprotein PAL (*H375_7520*), thymidylate kinase (*H375_8480*), and lipopolysaccharide ABC transporter (*H375_1050*) reveal identical trends of changes in their expression as observed in our RNA-Seq (Figure 1). Importantly, although expression profiles of all four genes varied depending on the growth temperature of HMECs (37°C versus 34°C), host niche (HMECs versus AAE2) also had a profound impact on the abundance of these transcripts.

The field of bacterial small RNAs has been evolving quite rapidly over the past few years, yet the potential contributions of rickettsial sRNAs during pathogen-host and pathogen-vector interactions remain poorly understood. To the best of our knowledge, this is the first differential transcriptomics study of *R. prowazekii* in human and vector cells as the host. Precise identification and selective use of transcriptional start sites for rickettsial genes in a particular host furthers our understanding of genome organization and plasticity and discovery of hitherto unknown highly abundant sRNAs unique to arthropod host cells poses new queries related to their functions in *R. prowazekii*. Transcriptome analyses of other pathogenic rickettsiae, which exploit ticks as their predominant natural vectors, are currently ongoing and should reveal further insights into the roles of ribo-regulatory mechanisms among members of *Rickettsiales*.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding

This work was supported by National Institute of Allergy and Infectious Diseases at the National Institutes of Health [grant 5R21 AI115231-02], and in part by a pilot project grant from the Institute for Human Infections and Immunity, the James McLaughlin Fellowship Program, and institutional support funds from the University of Texas Medical Branch.

We thank Drs. Thomas Wood and Steve Widen for their help with library preparation and RNA sequencing.

Abbreviations

AAE2	*Amblyomma americanum* cells
CDS	coding sequence
gltA	citrate synthase gene
HMECs	human microvascular endothelial cells
MEV	mean expression value
***Rp*_sR**	*Rickettsia prowazekii* small ribonucleic acid
RPKM	reads per kilobase per million
TEX	terminator 5′-phosphate-dependent exonuclease
TPM	transcripts per kilobase million
TSS	transcription start site

References

Barnard A, Wolfe A, Busby S. Regulation at complex bacterial promoters: how bacteria use different promoter organizations to produce different regulatory outcomes. Curr Opin Microbiol. 2004; 7(2): 102–108. [PubMed: 15063844]

Bechah Y, Capo C, Mege JL, Raoult D. Epidemic typhus. Lancet Infect Dis. 2008a; 8(7):417–426. [PubMed: 18582834]

Bechah Y, Capo C, Mege JL, Raoult D. Rickettsial diseases: from *Rickettsia*-arthropod relationships to pathophysiology and animal models. Future Microbiol. 2008b; 3(2):223–236. [PubMed: 18366341]

Bozeman FM, Masiello SA, Williams MS, Elisberg BL. Epidemic typhus rickettsiae isolated from flying squirrels. Nature. 1975; 255(5509):545–547. [PubMed: 806809]

Busch A, Richter AS, Backofen R. IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. Bioinformatics. 2008; 24:2849–2856. [PubMed: 18940824]

Cai J, Pang H, Wood DO, Winkler HH. The citrate synthase-encoding gene of *Rickettsia prowazekii* is controlled by two promoters. Gene. 1995; 163(1):115–119. [PubMed: 7557459]

Chan YG, Cardwell MM, Hermanas TM, Uchiyama T, Martinez JJ. Rickettsial outer-membrane protein B (rOmpB) mediates bacterial invasion through Ku70 in an actin, c-Cbl, clathrin and caveolin 2-dependent manner. Cell Microbiol. 2009; 11(4):629–644. [PubMed: 19134120]

Dreher-Lesnick SM, Ceraul SM, Rahman MS, Azad AF. Genome-wide screen for temperature-regulated genes of the obligate intracellular bacterium, *Rickettsia typhi*. BMC Microbiol. 2008; 8:61. [PubMed: 18412961]

Dugar G, Herbig A, Förstner KU, Heidrich N, Reinhardt R, Nieselt K, Sharma CM. High-resolution transcriptome maps reveal strain-specific regulatory features of multiple *Campylobacter jejuni* isolates. PLoS Genet. 2013; 9(5):e1003495. [PubMed: 23696746]

Ellison DW, Clark TR, Sturdevant DE, Virtaneva K, Hackstadt T. Limited transcriptional responses of *Rickettsia rickettsii* exposed to environmental stimuli. PLoS One. 2009; 4(5):e5612. [PubMed: 19440298]

Galletti MFBM, Fujita A, Nishiyama MY Jr, Malossi CD, Pinter A, Soares JF, Daffre S, Labruna MB, Fogaca AC. Natural blood feeding and temperature shift modulate the global transcriptional profile of *Rickettsia rickettsii* infecting its tick vector. PLoS One. 2013; 8(10):e77388. [PubMed: 24155949]

Gillespie JJ, Ammerman NC, Dreher-Lesnick SM, Rahman MS, Worley MJ, Setubal JC, Sobral BS, Azad AF. An anomalous type IV secretion system in *Rickettsia* is evolutionarily conserved. PloS one. 2009; 4(3):e4833. [PubMed: 19279686]

Gillespie JJ, Kaur SJ, Rahman MS, Rennoll-Bankert K, Sears KT, Beier-Sexton M, Azad AF. Secretome of obligate intracellular *Rickettsia*. FEMS Microbiol Rev. 2015; 39(1):47–80. [PubMed: 25168200]

Gillespie JJ, Phan IQ, Driscoll TP, Guillotte ML, Lehman SS, Rennoll-Bankert KE, Subramanian S, Beier-Sexton M, Myler PJ, Rahman MS, Azad AF. The *Rickettsia* type IV secretion system: unrealized complexity mired by gene family expansion. Pathogens and disease. 2016; 74(6)

Gottesman S, Storz G. Bacterial small RNA regulators: versatile roles and rapidly evolving variations. Cold Spring Harb Perspect Biol. 2011; 3(12):a003798–a003798. [PubMed: 20980440]

Khandige S, Kronborg T, Uhlin BE, Møller-Jensen J. sRNA-mediated regulation of P-fimbriae phase variation in uropathogenic *Escherichia coli*. PLoS Path. 2015; 11(8):e1005109.

Labruna MB, Whitworth T, Horta MC, Bouyer DH, McBride JW, Pinter A, Popov V, Gennari SM, Walker DH. *Rickettsia* species infecting *Amblyomma cooperi* ticks from an area in the state of Sao Paulo, Brazil, where Brazilian spotted fever is endemic. J Clin Microbiol. 2004; 42(1):90–98. [PubMed: 14715737]

Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012; 9(4):357–359. [PubMed: 22388286]

Lasa I, Toledo-Arana A, Dobin A, Villanueva M, de los Mozos IR, Vergara-Irigaray M, Segura V, Fagegaltier D, Penadés JR, Valle J, Solano C, Gingeras TR. Genome-wide antisense transcription

drives mRNA processing in bacteria. Proc Natl Acad Sci U S A. 2011; 108(50):20172–20177. [PubMed: 22123973]

Lasa I, Toledo-Arana A, Gingeras TR. An effort to make sense of antisense transcription in bacteria. RNA Biol. 2012; 9(8):1039–1044. [PubMed: 22858676]

Li B, Ruotti V, Stewart RM, Thomson JA, Dewey CN. RNA-Seq gene expression estimation with read mapping uncertainty. Bioinformatics. 2010; 26(4):493–500. [PubMed: 20022975]

Li J, Qi L, Guo Y, Yue L, Li Y, Ge W, Wu J, Shi W, Dong X. Global mapping transcriptional start sites revealed both transcriptional and post-transcriptional regulation of cold adaptation in the methanogenic archaeon Methanolobus psychrophilus. Scientific reports. 2015; 5:9209. [PubMed: 25784521]

Liu JM, Camilli A. A broadening world of bacterial small RNAs. Curr Opin Microbiol. 2010; 13(1): 18–23. [PubMed: 20022798]

McGrath PT, Lee H, Zhang L, Iniesta AA, Hottes AK, Tan MH, Hillson NJ, Hu P, Shapiro L, McAdams HH. High-throughput identification of transcription start sites, conserved promoter motifs and predicted regulons. Nat Biotechnol. 2007; 25(5):584–592. [PubMed: 17401361]

Medina-Sanchez A, Bouyer DH, Alcantara-Rodriguez V, Mafra C, Zavala-Castro J, Whitworth T, Popov VL, Fernandez-Salas I, Walker DH. Detection of a typhus group *Rickettsia* in *Amblyomma* ticks in the state of Nuevo Leon, Mexico. Ann N Y Acad Sci. 2005; 1063:327–332. [PubMed: 16481535]

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nature methods. 2008; 5(7):621–628. [PubMed: 18516045]

Munderloh UG, Kurtti TJ. Formulation of medium for tick cell culture. Exp Appl Acarol. 1989; 7(3): 219–229. [PubMed: 2766897]

Narra HP, Schroeder CL, Sahni A, Rojas M, Khanipov K, Fofanov Y, Sahni SK. Small Regulatory RNAs of *Rickettsia conorii*. Scientific reports. 2016; 6:36728. [PubMed: 27834404]

Ohashi N, Nashimoto H, Ikeda H, Tamura A. Diversity of immunodominant 56-kDa type-specific antigen (TSA) of *Rickettsia tsutsugamushi*. Sequence and comparative analyses of the genes encoding TSA homologues from four antigenic variants. J Biol Chem. 1992; 267(18):12728–12735. [PubMed: 1618776]

Ooi WF, Ong C, Nandi T, Kreisberg JF, Chua HH, Sun G, Chen Y, Mueller C, Conejero L, Eshaghi M, Ang RML, Liu J, Sobral BW, Korbsrisate S, Gan YH, Titball RW, Bancroft GJ, Valade E, Tan P. The condition-dependent transcriptional landscape of *Burkholderia pseudomallei*. PLoS Genet. 2013; 9(9):e1003795. [PubMed: 24068961]

Östberg Y, Bunikis I, Bergström S, Johansson J. The etiological agent of Lyme Disease, *Borrelia burgdorferi*, appears to contain only a few small RNA molecules. J Bacteriol. 2004; 186(24):8472–8477. [PubMed: 15576797]

Philip CB, Hoogstraal H, Reiss-Gutfreund R, Clifford CM. Evidence of rickettsial disease agents in ticks from Ethiopian cattle. Bull WHO. 1966; 35(2):127–131. [PubMed: 5296996]

Raoult D, Woodward T, Dumler JS. The history of epidemic typhus. Infect Dis Clin North Am. 2004; 18(1):127–140. [PubMed: 15081509]

Rydkina E, Silverman DJ, Sahni SK. Activation of p38 stress-activated protein kinase during *Rickettsia rickettsii* infection of human endothelial cells: role in the induction of chemokine response. Cell Microbiol. 2005; 7:1519–1530. [PubMed: 16153249]

Schroeder CL, Narra HP, Rojas M, Sahni A, Patel J, Khanipov K, Wood TG, Fofanov Y, Sahni SK. Bacterial small RNAs in the Genus *Rickettsia*. BMC genomics. 2015; 16(1):1075. [PubMed: 26679185]

Schroeder CLC, Narra HP, Sahni A, Rojas M, Khanipov K, Patel J, Shah R, Fofanov Y, Sahni SK. Identification and characterization of novel small RNAs in *Rickettsia prowazekii*. Front Microbiol. 2016; 7(859)

Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, Chabas S, Reiche K, Hackermuller J, Reinhardt R, Stadler PF, Vogel J. The primary transcriptome of the major human pathogen *Helicobacter pylori*. Nature. 2010; 464(7286):250–255. [PubMed: 20164839]

Sharma CM, Vogel J. Differential RNA-seq: the approach behind and the biological insight gained. Curr Opin Microbiol. 2014; 19:97–105. [PubMed: 25024085]

Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C_T method. Nature Protocols. 2008; 3:1101–1108. [PubMed: 18546601]

Stubben CJ, Micheva-Viteva SN, Shou Y, Buddenborg SK, Dunbar JM, Hong-Geller E. Differential expression of small RNAs from *Burkholderia thailandensis* in response to varying environmental and stress conditions. BMC Genomics. 2014; 15(1):385. [PubMed: 24884623]

Uchiyama T. Tropism and pathogenicity of rickettsiae. Front Microbiol. 2012; 3:230. [PubMed: 22737150]

Van Houten B, Kad N. Investigation of bacterial nucleotide excision repair using single-molecule techniques. DNA Repair. 2014; 20:41–48. [PubMed: 24472181]

Wade JT. Where to begin? Mapping transcription start sites genome-wide in *Escherichia coli*. J Bacteriol. 2015; 197(1):4–6. [PubMed: 25331438]

Wagner GP, Kin K, Lynch VJ. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theory in Biosciences. 2012; 131(4):281–285. [PubMed: 22872506]

Walker DH, Ismail N. Emerging and re-emerging rickettsioses: endothelial cell infection and early disease events. Nat Rev Microbiol. 2008; 6(5):375–386. [PubMed: 18414502]

Waters LS, Storz G. Regulatory RNAs in bacteria. Cell. 2009; 136(4):615–628. [PubMed: 19239884]

Westermann AJ, Gorski SA, Vogel J. Dual RNA-seq of pathogen and host. Nat Rev Microbiol. 2012; 10(9):618–630. [PubMed: 22890146]

Woolfit M, Algama M, Keith JM, McGraw EA, Popovici J. Discovery of putative small non-coding RNAs from the obligate intracellular bacterium *Wolbachia pipientis*. PLoS One. 2015; 10(3):e0118595. [PubMed: 25739023]

Wozniak CE, Chevance FFV, Hughes KT. Multiple promoters contribute to swarming and the coordination of transcription with flagellar assembly in *Salmonella*. J Bacteriol. 2010; 192(18): 4752–4762. [PubMed: 20639318]

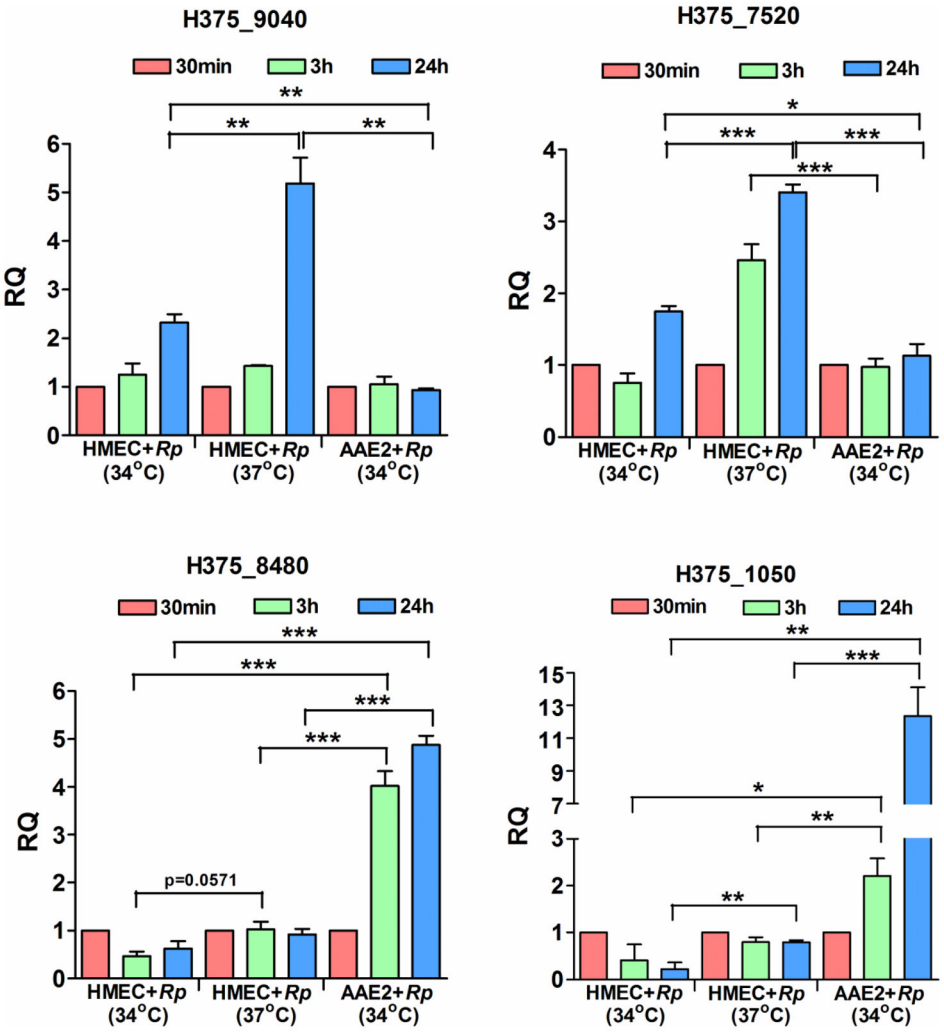

Figure 1. Quantitative PCR based validation of *R. prowazekii* genes differentially expressed during host and tick cell infection *in vitro*

Confluent monolayers of human microvascular endothelial cells (HMECs) grown at 37°C or 34°C or *Amblyomma americanum* cells (AAE2) grown at 34°C were infected with *R. prowazekii* for 0.5h, 3h, and 24h. Total RNA was extracted using Trizol, treated with DNaseI, and reverse transcribed for RT-PCR (n≥3). Expression profile of four genes namely, *H375_9040*, *H375_7520*, *H375_8480* and *H375_1050* was quantified at early (3h) and late (24h) stages of infection of HMECs and AAE2 cells using gene specific primers and 16S rRNA as endogenous control. The data is calculated using expression levels at 0.5h as the baseline and presented as mean±SEM. Asterisks indicate *p<0.05, **p<0.01 and ***p<0.001.

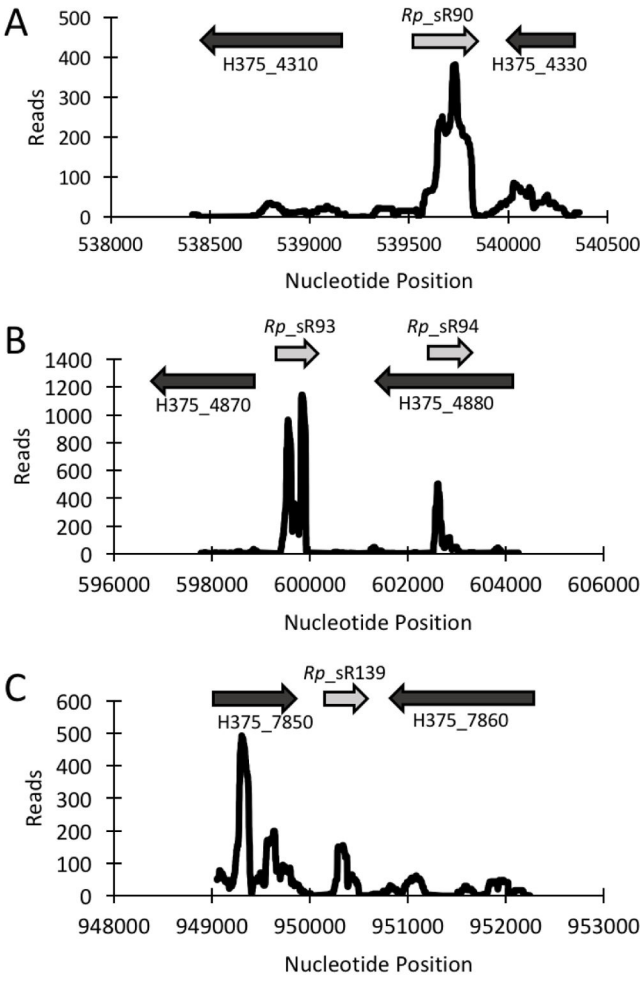

Figure 2. Expression profile of novel t*rans-Acting* AAE2 specific *Rp*_sRs

Shown are the coverage plots for selected *trans*-acting *Rp*_sRs observed during the infection of AAE2 cells. Nucleotide positions within the genome are indicated on X-axis and the Y-axis displays the number of reads for that particular nucleotide position. The light grey arrow represents the small RNA. The dark grey arrows represent the orientation of upstream and downstream ORFs, respectively.

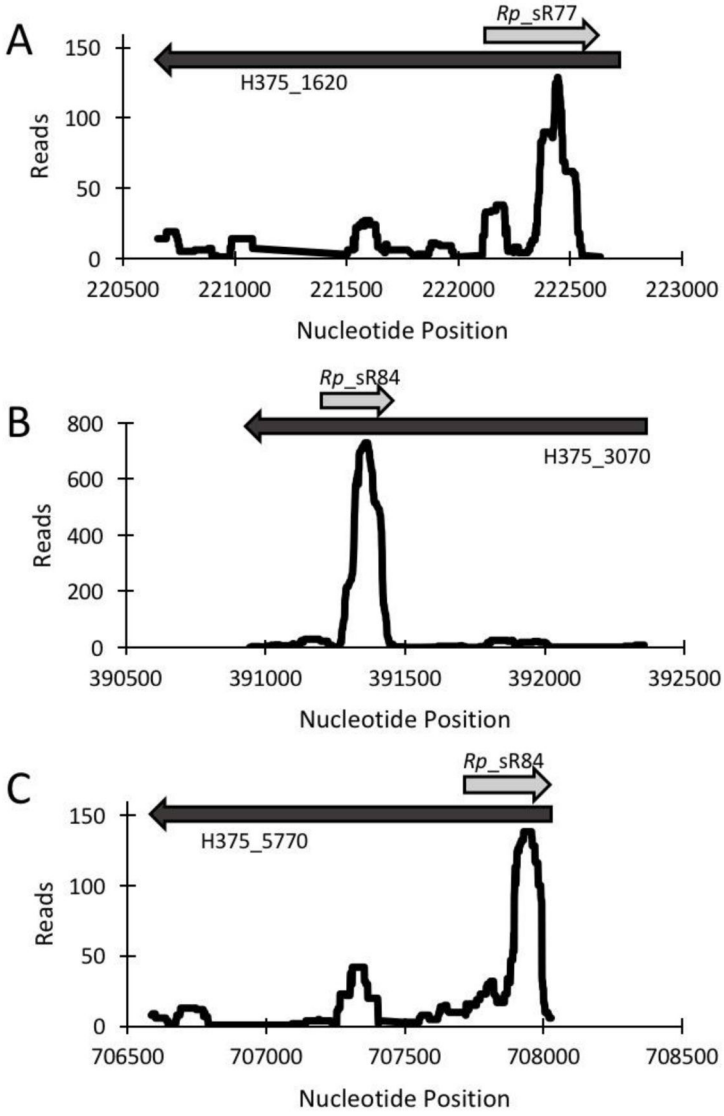

Figure 3. Expression profile of novel *Cis-Acting* AAE2 Specific *Rp*_sRs

Shown are the coverage plots for selected *cis*-acting *Rp*_sRs observed during the infection of AAE2 cells. Nucleotide positions within the genome are indicated on X-axis and the Y-axis displays the number of reads for that particular nucleotide position. The light grey arrow represents the small RNA. The dark grey arrows represent the orientation of the respective ORF.

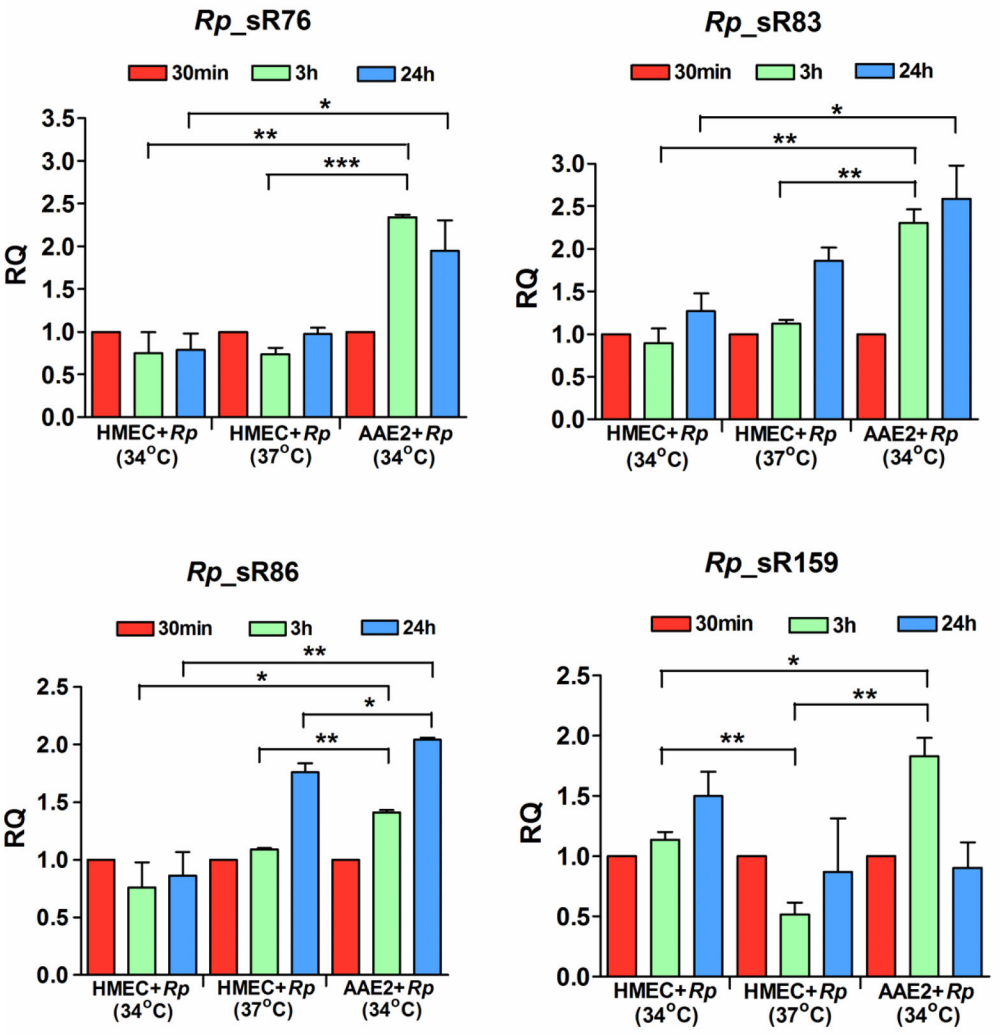

Figure 4. qPCR of novel differentially expressed *Rp*_sRs

Confluent monolayers of human microvascular endothelial cells (HMECs) cultured at 37°C or 34°C or *Amblyomma americanum* cells (AAE2) cultured at 34°C were infected with *R. prowazekii* for 0.5h, 3h, and 24h. Total RNA was extracted using Trizol, DNaseI treated, and reverse transcribed for RT-PCR (n≥3). AAE2 and HMEC expression were baselined to 0.5h and normalized to 16S rRNA. The data is presented as mean±SEM for four trans-acting *Rp*_sRs, namely, *Rp*_sR76, *Rp*_sR83, *Rp*_sR86, and *Rp*_sR159. Asterisks indicate *p<0.05, **p<0.01 and ***p<0.001.

Table 1

List of *R. prowazekii* ORFs upregulated during the infection of HMECs when compared to AAE2 cells *in vitro*.

Gene	TPM	Annotation
H375_9040	22.93	Heat shock protein 60 family co-chaperone GroES
H375_7520	12.67	Tol-Pal system peptidoglycan-associated lipoprotein PAL
H375_6910	11.08	Putative adhesion (homolog of Adr1)
H375_3980	9.81	Threonyl-tRNA synthetase
H375_450	8.23	Cell division protein MraZ
H375_7750	7.84	Twin-arginine translocation protein TatA
H375_6090	6.64	ATP synthase F0 sector subunit c (EC 3.6.3.14)
H375_0400	5.66	Uncharacterized protein RT0563
H375_3470	5.66	rickettsial conserved
H375_5060	3.92	HflC protein
H375_9050	3.79	Heat shock protein 60 family chaperone GroEL
H375_4550	3.77	DNA-binding protein HU
H375_1770	3.76	Succinyl-CoA ligase [ADP-forming] alpha chain
H375_8520	3.62	2-methoxy-6-polyprenyl-1,4-benzoquinol methylase
H375_4120	3.45	Peptide deformylase
H375_4630	3.45	rickettsial conserved
H375_5960	3.45	Acetoacetyl-CoA reductase
H375_1100	3.02	SSU ribosomal protein S15p
H375_3430	4.66	Small heat shock protein C1
H375_9200	4.61	SSU ribosomal protein S21p
H375_6920	4.58	hypothetical protein
H375_9250	4.53	hypothetical protein
H375_8350	4.39	rickettsial conserved
H375_1700	4.32	rickettsial conserved
H375_6600	4.02	hypothetical protein
H375_7550	3.28	Rod shape-determining protein MreC
H375_8110	3.15	hypothetical protein
H375_2030	3.02	Cytochrome c oxidase polypeptide II
H375_3520	3.02	Outer membrane protein H precursor
H375_7430	3.02	rickettsial conserved
H375_7600	3.02	Acyl carrier protein
H375_7730	3.02	LSU ribosomal protein L21p
VBIRicPro269054_0891	3.02	hypothetical protein
H375_1740	3.02	Ribosome-binding factor A

Table 2

List of *R. prowazekii* ORFs upregulated during the infection of AAE2 cells when compared to HMECs *in vitro*.

Gene	TPM	Annotation
H375_8480	28.84	Thymidylate kinase
H375_1050	23.2	Lipopolysaccharide ABC transporter
H375_2780	22.87	dTDP-4-dehydrorhamnose reductase
H375_6110	22.21	ATP synthase F0 sector subunit b
H375_2300	18.23	3-deoxy-manno-octulosonate cytidylyltransferase
H375_7500	15.25	LSU ribosomal protein L32p
H375_7840	14.91	Metal-dependent hydrolase YbeY
H375_1210	12.76	rickettsial conserved
H375_4640	12.37	hypothetical protein
H375_9060	12.35	guanosine-3,5-bis(diphosphate) 3-pyrophosphohydrolase SpoTc
H375_2110	12.26	Signal peptide peptidase SppA, 36K type
H375_0670	11.93	Citrate synthase
H375_8910	11.77	LSU ribosomal protein L30p
H375_1070	10.94	Lipopolysaccharide-assembly protein LptC
H375_0860	10.77	Fumarate hydratase class II
H375_5480	10.27	hypothetical protein
H375_7110	9.69	YrbA protein
H375_1170	9.61	Multicopper polyphenol oxidase
H375_0490	9.53	LSU ribosomal protein L10p
H375_4210	9.28	rickettsial conserved
H375_3780	8.95	Inositol-1-monophosphatase
H375_6030	8.78	DNA recombination and repair protein RecF
H375_5500	8.12	hypothetical protein
H375_0500	8.07	Succinate dehydrogenase hydrophobic membrane anchor protein
H375_1840	7.79	Undecaprenyl diphosphate synthase
H375_6230	7.79	3-hydroxyacyl-[acyl-carrier-protein] dehydratase
H375_6070	7.73	ATP synthase protein I
H375_7490	7.62	Outer membrane lipoprotein OmlA
H375_4230	7.46	Universal stress protein family 3
H375_6390	7.46	Octanoate-[acyl-carrier-protein]-protein-N-octanoyltransferase
H375_7130	7.29	Putative glutathione-regulated potassium-efflux system protein KefB
H375_8740	6.96	LSU ribosomal protein L4p
H375_2600	4.64	Exodeoxyribonuclease VII small subunit
H375_0780	4.47	hemolysin C
H375_5450	4.47	SSU ribosomal protein S2p (SAe)
H375_2690	4.4	RND efflux system
H375_0850	4.39	Outer membrane protein Imp
H375_6020	4.38	hypothetical protein

Gene	TPM	Annotation
H375_3220	4.37	Guanosine pentaphosphate phosphohydrolase
H375_4250	4.27	Multimodular transpeptidase-transglycosylase
H375_8940	4.14	Adenylate kinase
H375_1780	4.14	Succinyl-CoA ligase
H375_2760	4.08	UDP-N-acetyl-L-fucosamine synthase
H375_7700	4.02	hypothetical protein
H375_4520	4.01	Signal recognition particle protein Ffh
H375_5310	4	LSU ribosomal protein L28p
H375_5490	3.98	hypothetical protein
H375_6300	3.98	Uroporphyrinogen III decarboxylase
H375_2020	3.95	hypothetical protein
H375_2480	3.93	hypothetical protein
H375_6800	3.92	hypothetical protein
H375_2010	3.91	Lipoprotein signal peptidase
H375_5900	3.87	LSU ribosomal protein L9p
H375_0960	3.81	carbonic anhydrase, family 3
H375_7970	3.79	NADH:ubiquinone oxidoreductase 17.2 kD subunit
H375_5050	3.76	HtrA protease/chaperone protein
H375_3930	3.74	hypothetical protein
H375_2360	3.73	RalF protein
H375_2200	3.72	rare lipoprotein A precursor
H375_8000	3.71	tRNA-guanine transglycosylase
H375_1890	3.71	hypothetical protein
H375_7280	3.65	NADH-ubiquinone oxidoreductase chain I
H375_4740	3.65	Aspartyl-tRNA(Asn) amidotransferase subunit C
H375_4180	3.65	Chaperone protein HscB
H375_7380	6.96	hypothetical protein
H375_6130	6.52	cell surface antigen
H375_0060	6.46	RNA binding methyltransferase FtsJ like
H375_0950	6.46	Ribosome hibernation protein YhbH
H375_6190	6.46	hypothetical protein
H375_1730	6.37	rickettsial conserved
H375_8460	6.3	4-hydroxybenzoate polyprenyltransferase
H375_4360	6.3	serine protease, HtrA/DegQ/DegS family
H375_6010	5.83	Cytochrome oxidase biogenesis protein
H375_5370	5.83	Phospholipid ABC transporter shuttle protein MlaC
H375_0020	5.8	hypothetical protein
H375_6040	5.72	Methyltransferase
H375_3940	5.66	Type I secretion outer membrane protein, TolC precursor
H375_3010	5.63	hypothetical protein
H375_2720	5.5	Membrane protein, putative
H375_1830	5.47	Osmolarity sensory histidine kinase EnvZ

Gene	TPM	Annotation
H375_3830	5.44	SSU ribosomal protein S9p (S16e)
H375_0160	5.44	ATP-dependent protease La (EC 3.4.21.53) Type I
H375_3020	5.37	TolA protein
H375_1940	5.36	Diaminopimelate epimerase
H375_2490	5.3	rickettsial conserved
H375_1920	5.3	Phenylalanyl-tRNA synthetase alpha chain
H375_2550	5	NADH-ubiquinone oxidoreductase chain C
H375_9240	4.97	hypothetical protein
H375_7670	4.97	hypothetical protein
H375_5710	4.97	Chromosome (plasmid) partitioning protein ParB/Stage 0 sporulation protein J
H375_5660	4.91	dNTP triphosphohydrolase, broad substrate specificity
H375_2140	4.81	rickettsial conserved
H375_1870	4.79	Ribonuclease D related protein
H375_3390	4.72	rickettsial conserved
H375_0470	4.72	Ribosome recycling factor
H375_7940	4.64	UDP-2,3-diacylglucosamine pyrophosphatase
H375_2710	3.6	minor teichoic acids biosynthesis protein ggaB
H375_4420	3.56	Thermostable carboxypeptidase 1
H375_2700	3.55	hypothetical protein
H375_6310	3.42	Ferrochelatase
H375_3910	3.42	Topoisomerase IV subunit B
H375_7170	3.4	Ferroxidase
H375_7530	3.37	hypothetical protein
H375_1990	3.37	UDP-N-acetylmuramoylalanine
H375_8240	3.35	Integration host factor alpha subunit
H375_8610	3.31	rickettsial conserved
H375_0140	3.31	Rhodanese-related sulfurtransferase
H375_1010	3.31	rickettsial conserved
H375_4140	3.28	Multidrug resistance protein Atm1
H375_3250	3.27	VirB10
H375_7190	3.26	ATP synthase delta chain
H375_3180	3.26	rickettsial conserved
H375_0600	3.2	Bacterial ribosome SSU maturation protein RimP
H375_3130	3.2	Thymidylate synthase ThyX
H375_6290	3.2	Phosphoenolpyruvate synthase regulatory protein
H375_5890	3.2	tRNA(Ile)-lysidine synthetase
H375_5040	3.2	Rhodanese domain protein
H375_8080	3.15	Mobile element protein
H375_4990	3.15	SSU ribosomal protein S12p (S23e)
H375_5840	3.15	Small-conductance mechanosensitive channel
H375_7680	3.13	hypothetical protein
H375_0540	3.12	2-octaprenyl-6-methoxyphenol hydroxylase

Gene	TPM	Annotation
H375_4400	3.08	lipoprotein ComL
H375_4050	3.03	rickettsial conserved
H375_7420	3.01	rickettsial conserved
H375_0520	3.01	VirB6

Table 3

List of *R. prowazekii* ORFs exhibiting different transcription start sites during their *in vitro* expression in HMECs and AAE2 cells

Locus Tag	HMEC-T	AAE2-T	Strand	Annotation
H375_0890	119283	119025	+	Transcription termination factor Rho
H375_0950	128486	128393	+	Ribosome hibernation protein YhbH
H375_2140	285600	285557	+	Rickettsial conserved
H375_3140	400166	400204	−	Membrane-bound metallopeptidase
H375_3150	400514	400453	+	Ribose 5-phosphate isomerase B
H375_3390	427957	427919	+	Hypothetical protein
H375_3530	443893	443878	+	GTP-binding protein TypA
H375_3540	447237	447188	−	Pyruvate dehydrogenase E1 subunit
H375_3630	457043	457059	+	UDP-3-O-[3-hydroxymyristoyl] N-Acetylglucosamine deacetylase
H375_3640	458837	458779	+	Membrane c-type cytochrome cy
H375_3710	466907	466938	−	Hypothetical protein
H375_3800	472515	472467	+	Pole remodelling regulatory diguanylate cyclase
H375_3940	488473	488445	−	Type I secretion outer membrane protein, TolC
H375_4120	513479	513553	−	Peptide deformylase
H375_4800	591148	591116	+	Integral membrane protein
H375_5400	666077	666043	+	Aspartate aminotransferase
H375_8580	1041290	1041233	+	Outer membrane protein Imp
H375_9020	1076523	1076555	−	Heat shock protein GrpE