Detection and quantification of intergenic transcription in *Mycoplasma hyopneumoniae*

Stuart W. Gardner\(^1,2\) and F. Chris Minion\(^1\)

\(^1\)Department of Veterinary Microbiology and Preventive Medicine, Interdepartmental Microbiology Program, Iowa State University, Ames, IA 50011, USA

\(^2\)Department of Statistics, Iowa State University, Ames, IA 50011, USA

Mycoplasmas are thought to control gene expression through simple mechanisms. The switching mechanisms needed to regulate transcription during significant environmental shifts do not seem to be required for these host-adapted organisms. *Mycoplasma hyopneumoniae*, a swine respiratory pathogen, undergoes differential gene expression, but as for all mycoplasmas, the mechanisms involved are still unknown. Since mycoplasmas contain only a single sigma factor and few regulator-type proteins, it is likely that other mechanisms control gene regulation, possibly involving intergenic (IG) regions. To study this further, we investigated whether IG regions are transcribed in *M. hyopneumoniae*, and measured transcription levels across five specific regions. Microarrays were constructed with probes covering 343 IG regions of the *M. hyopneumoniae* genome, and RNA isolated from laboratory-grown cells was used to interrogate the arrays. Transcriptional signals were identified in 321 (93.6\%) of the IG regions. Five large (>500 bp) IG regions were chosen for further analysis by qRT-PCR by designing primer sets whose products reside in flanking ORFs, bridge flanking ORFs and the IG region, or reside solely within the IG region. The results indicate that no single transcriptional start site can account for transcriptional activity within IG regions. Transcription can end abruptly at the end of an ORF, but this does not seem to occur at high frequency. Rather, transcription continues past the end of the ORF, with RNA polymerase gradually releasing the template. Transcription can also be initiated within IG regions in the absence of accepted promoter-like sequences.

INTRODUCTION

Mycoplasmas are members of the class *Mollicutes*, cell-wall-less bacteria that infect a wide variety of plants and animals (including humans). Though related taxonomically, members of the *Mollicutes* differ significantly in their habitats, their growth requirements and their overall structure. They are believed to have evolved from more complex organisms through degenerative mechanisms with the concomitant loss of genetic material (Neimark, 1986); all mycoplasmas have small genomes with a limited number of genes, resulting in a lack of biosynthetic pathways (Hutchison & Montague, 2002; Pollack et al., 1997). Thus, they must obtain amino acids, purines, pyrimidines and membrane components from their growth environment. For the members of the family *Mycoplastataceae*, the animal pathogens, this occurs primarily at the host mucosal surface, where damage to host cells occurs through different mechanisms (Razin & Herrmann, 2002).

Mycoplasmas are found in association with hosts and not in a variety of external environments like many other pathogenic bacteria. Thus, the question arises as to whether mycoplasmas need to regulate genes if their environments are restricted to host mucosal surfaces, and if they do, then by what mechanisms. It has been suggested that with the loss of biosynthetic pathways, the remaining genes that are responsible for the basic processes of the cell, i.e. DNA replication, transcription and translation, are thought to be constitutively expressed, and in constant environments, the organism supposedly has little need for sophisticated genetic control mechanisms (Muto & Ushida, 2002). Our knowledge of gene expression in mycoplasmas is rudimentary, with few studies that actually address the mechanisms involved (Benders et al., 2005; Chang et al., 2008; Güell et al., 2009; Lluch-Senar et al., 2007; Weiner et al., 2000, 2003). Transcription is essential to the basic process of gene expression, but surprisingly, little information has been obtained experimentally defining the transcription process in mycoplasmas (Muto & Ushida, 2002). The heat-shock response has been studied in a

Abbreviations: IG, intergenic; qRT-PCR, quantitative real-time-PCR.

The microarray descriptions and the microarray data can be accessed through the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo) under accession number GSE17893.

A set of supplementary methods, four supplementary figures and five supplementary tables are available with the online version of this paper.
limited number of mycoplasma species (Madsen et al., 2006b; Musatovova et al., 2006; Weiner et al., 2003), and it is clear that a transcriptional activator, HrcA, is involved in regulating transcription of two heat-shock genes (Chang et al., 2008). In addition, transcriptional start sites have been mapped for a limited number of genes (Hyman et al., 1988; Taschke & Herrmann, 1988; Taschke et al., 1986; Weiner et al., 2000), and promoter-probe vectors have been used in several different species (Janis et al., 2005; Knudtson & Minion, 1993, 1994; Lluch-Senar et al., 2007). These studies show that the –10 box of the Pribnow sequence is relatively conserved, but not the –35 sequence. This lack of progress has been partly due to the few genetic tools available for use with mycoplasmas and with some species, e.g. *Mycoplasma hyopneumoniae*, the total lack of such tools. The limited genome capacity suggests simple regulatory mechanisms, and experimental evidence indicates that mycoplasmas have few regulatory systems for several reasons: (1) only a single sigma factor has been identified in any of the complete mycoplasma genome sequences, (2) transcriptional activators have rarely been identified (Chang et al., 2008), and (3) the genes for a few repressor-like proteins have been identified in genome sequence studies (Fraser et al., 1995; Glass et al., 2000; Minion et al., 2004). The global assessment of transcription has been reported using arrays (Madsen et al., 2006a, b, 2008; Oneal et al., 2008; Schafer et al., 2007; Weiner et al., 2003), but these tools are limited because they only assess transcriptional changes in annotated gene sequences and do not address mechanisms of transcription control.

The subject of this study, *M. hyopneumoniae*, is associated with respiratory disease and reduced productivity in swine (Ross, 1992). It is the causative agent of porcine enzootic pneumonia, and when co-infections occur with a secondary (bacterial or viral) infection, the respiratory disease is more severe and has been designated porcine respiratory disease complex (Ross, 1992). Even though the virulence of *M. hyopneumoniae* appears to be low, substantial economic losses occur as a consequence of reduced average daily weight gain and efficiency of feed utilization, prophylactic and therapeutic costs, and mortality (Pointon et al., 1985; Ross, 1992). The virulence factors of *M. hyopneumoniae* are largely unknown, and to better understand the mechanisms involved, we are studying transcription in *M. hyopneumoniae* both in vitro and in vivo (Adams et al., 2005; Madsen et al., 2006a, b, 2008; Oneal et al., 2008; Schafer et al., 2007). Recent studies using microarrays to examine global transcriptional changes clearly show that *M. hyopneumoniae* can regulate its genes in response to environmental changes (Madsen et al., 2006a, b, 2008; Oneal et al., 2008; Schafer et al., 2007). These studies give a snapshot of transcription across the genome in the annotated genes, but they cannot reveal basic features of transcription, i.e. start and stop sequences and mechanisms of regulation.

Many mycoplasma genes lack defined, typical or canonical transcriptional stop sequences; rho-dependent termination does not seem to occur in mycoplasmas (Razin et al., 1998) and rho-independent sequences cannot often be found. To better understand transcription termination in mycoplasmas and as a first step in defining potential regulatory RNAs that may serve a role in gene regulation (Gottesman, 2005; Szymanski et al., 2003; Waters & Storz, 2009), we sought to determine if intergenic (IG) regions were transcribed in *M. hyopneumoniae*. The genome of *M. hyopneumoniae* consists of 698 protein-coding ORFs and 35 rRNA genes, and due to overlapping and juxtaposed ORFs there are 573 IG regions (Minion et al., 2004). To complicate the study, the overall G+C content of the genome is 28.6 mol%, with ORFs having 29.6 mol% and IG regions 18.7 mol% G+C. Surprisingly, the results of this study indicate that transcripts specific to virtually all of the 343 IG regions longer than 50 bases can be identified in *M. hyopneumoniae* RNA preparations. In a more defined quantitative analysis of five IG regions, it was clearly shown that internal initiation of transcription in IG regions occurs and probably does so at high frequency, suggesting that these RNAs may play an important role in the physiology of the mycoplasma, such as acting as regulatory RNAs. To our knowledge this is the first study to assess transcription across IG regions of the *M. hyopneumoniae* genome. A recent report by Güell et al. (2009) also demonstrated IG transcription in *Mycoplasma pneumoniae*, suggesting that transcripts arising from IG regions are common in mycoplasmas.

**METHODS**

**Mycoplasma strains and culture conditions.** Pathogenic *M. hyopneumoniae* strain 232, a derivative of strain 11, was used in this study (Mare & Switzer, 1965). Cultures were passed fewer than 15 times in vitro in Fries medium (Fries, 1975). For the microarray and IG transcriptional studies, three 250 ml flasks containing 125 ml cultures were grown at 37 °C to early exponential phase, as determined by medium colour change and optical density. Cells were pelleted by centrifugation at 24,000 g, 500 µl RNAlater (Ambion) was added, and pellets were stored at –70 °C until the total RNA was isolated.

For the heat-shock study, six cultures of 5.6 ml were grown to exponential phase as described previously (Madsen et al., 2006b). Three cultures were randomly chosen and then shifted to 42 °C for 2 h while the other three cultures were maintained at 37 °C for these 2 h. Cells were then pelleted and stored as described above until total RNA isolation.

**RNA isolation.** RNA was isolated from frozen cell pellets using Ambion’s Ribopure kit according to the manufacturer’s protocol. Each sample was treated with 4 U DNase at 37 °C for 30 min followed by purification using a Millipore Ultracell YM-30 Microcon Centrifugal Filter unit. Quantification and purity checks of each sample were done using a Nanodrop ND-1000 spectrophotometer. RNA samples were confirmed to be DNA free via PCR and gel electrophoresis.

**Microarrays.** Two types of microarrays were constructed, one consisting of PCR products and the other of synthesized oligonucleotides. Intergenic regions were defined as non-coding DNA sequences between annotated genes in the *M. hyopneumoniae* genome sequence (Minion et al., 2004). Only IG regions larger than 49 bp were considered for analysis, and those greater than 124 bp were considered for PCR-based probes. Oligonucleotides were used for the smaller IG regions and for sequences that failed in PCRs to complete the coverage.
(defined as >50% of the IG region). PCR primers were designed for products of 125–450 bp using the Primer3 online software program (Rozen & Skaltsky, 2000). PCRs were performed in 96-well plate format in a MJ Research Dyad thermocycler in 100 μl reactions using the following conditions: 1× MgCl₂-free PCR buffer (Invitrogen); 3.5 mM MgCl₂; 0.2 mM each of dATP, dTTP, dGTP and dCTP; 200 pmol of each primer; 2 units Taq polymerase (Invitrogen); and 100 ng M. hyopneumoniae strain 232 chromosomal DNA obtained via phenol/chloroform extraction and ethanol precipitation. Thermocycler conditions started with denaturation at 95 °C for 5 min followed by 30 cycles of denaturation at 94 °C for 1 min, annealing at 56 °C for 1 min, and elongation at 68 °C for 30 s, and ending with a final elongation at 68 °C for 10 min. Products were confirmed by 1.2% agarose gel electrophoresis and purified via Mo Bio’s UltraClean PCR Clean-Up kit according to the manufacturer’s instructions (Corning).

The purified PCR products were quantified by visual inspection, ending with a final elongation at 68 °C for 1 min followed by 30 cycles of denaturation at 94 °C, annealing at 56 °C, and elongation at 68 °C. The PCR products were resuspended to 100 pmol μl⁻¹ in Corning spotting buffer. Products suspended in spotting buffer were electrophoresed in 1.2% agarose gels to confirm purity and concentration of each product and its size. The PCR product microarrays were spotted to Corning UltraGAPS glass substrates in triplicate by the University of Iowa DNA facility.

Oligonucleotides were designed for some IG regions based on the inability to produce PCR products or the length of the region (49–124 bp). The strand chosen to be synthesized for the IG region probes was in the same orientation as the closest adjacent ORF. Oligonucleotides were synthesized by IDT and ranged in size between 50 and 60 bp. Oligonucleotides were resuspended to 100 pmol μl⁻¹ in nuclease-free water, and then 9 μl were vacuum dried and diluted to 30 μM in IDT Epoxide spotting buffer. These oligonucleotides were spotted to Corning Epoxide substrates in triplicate.

Microarray experimental design. To test for transcripts in IG regions, cDNA was generated from M. hyopneumoniae RNA using either random hexamers or a set of 129 hexamer oligonucleotide primers designed for M. hyopneumoniae mRNA (Madsen et al., 2006b). A single total RNA sample was split into eight 5 μg samples, and cDNA synthesis was conducted on four of these samples using random hexamers (IDT); for the other four samples, cDNA was generated using primers designed for M. hyopneumoniae as described by Madsen et al. (2006b). In each set of four, two samples were fluorescently labelled with Cy3 and the other two were labelled with Cy5 using an indirect labelling protocol (Madsen et al., 2006b). After labelling and cleanup, all eight samples were quantified on a Nanodrop ND-1000 spectrophotometer. For each array on the slide, 1 μg labelled cDNA from each type of primer generation (Cy3 or Cy5, random or specific primer set) was combined and hybridized to a PCR-featured array and an oligonucleotide-featured array. This resulted in two arrays of PCR probes hybridized with dye swaps and two arrays of oligonucleotide probes with dye swaps. A second RNA sample was split into four 5 μg samples of total RNA, and the hybridization process was repeated for the oligonucleotide probe array. The hybridization and wash protocols have been described for PCR (Madsen et al., 2006b) and oligonucleotide arrays [as per manufacturer’s instructions (Corning)].

Microarray data acquisition and analysis. Slides were scanned using a ScanArray Express laser scanner as described previously (Madsen et al., 2006b). The SoftWorRx Tracker analysis software package (Applied Precision) was used to quantify the images. Probes were declared to detect signal if their normalized intensity in both channels on at least one array was more than 6 SEM greater than the mean intensity of features containing spot buffer only.

Microarray data accession number. The microarray descriptions and the microarray data can be accessed through the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo) under accession number GSE17893.

qRT-PCR primer design. The IG sequence and 300 bases into adjacent ORFs were targeted for relative quantitative real-time-PCR (qRT-PCR) analysis. Primers were designed using the IDT Primer Quest tool. The goal was to design six primer pairs for each region with unique target sequences ranging in size from 80 to 200 bp. One target sequence each was designed to reside completely within each flanking ORF, one each to bridge the transition region from ORF to IG region, and two within the IG region itself. For clarity, the IG regions were designated according to the upstream ORF (numbered clockwise around the genome), independent of transcription direction. Additional primers for the heat-shock assay were similarly designed for the region upstream of dnaK. This region of the genome does not contain a large IG region; thus primers were designed such that a target sequence was located in each gene (uvrC, mhp071 and dnaK); and three target sequences bridged genomic/intergenic regions (uvrCig-070, mhp071ig-071 and ig-071iddnaK). Additionally, primers for a target sequence within the reference gene control (mhp381) used in the analysis were designed. Table 1 describes the primers used for the qRT-PCR analysis. All of the primers were tested using a M. hyopneumoniae chromosomal DNA template to ensure their effectiveness in producing a correct product as determined by size via agarose gel analysis.

qRT-PCR experimental design and data acquisition. A test of the sensitivity of the qRT-PCR assay was done using DNA from M. hyopneumoniae 232 diluted to 4 ng μl⁻¹. The Express One-Step SYBR GreenER kit (Invitrogen) was used for qRT-PCR with the Mx3005P QPCR System (Stratagene) and MxPro 4.1 software. Reaction volumes for each well totalled 15 μl and contained 0.895 μl water, 7.5 μl qPCR mix, 0.6 μl of each primer at 2.5 μM, 0.03 μl ROX reference dye, 5 μl template DNA or RNA at 4 ng μl⁻¹ and 0.375 μl SuperScript III. The DNA template was diluted 11 times as described by the PRECELX-Q software program (Gallup & Ackermann, 2006). These 11 dilutions and a negative control were run in a 96-well plate for the six target sequences in region ig-072. The thermal cycling conditions were a (mock) cDNA synthesis step at 50 °C for 30 min; RT deactivation at 95 °C for 10 min; 40 PCR cycles of 95 °C for 30 s, 56 °C for 30 s, and 68 °C for 30 s; and a melt-curve step of 95 °C for 10 min, 55 °C for 30 s, and 95 °C for 30 s. These thermal cycling conditions were used for all qRT-PCR assays throughout this study.

Test plates (Gallup & Ackermann, 2006) were run on five qRT-PCR regions (ig-072, ig-106, ig-282, ig-682, ig-684) for detection of inhibition and primer specificity of qRT-PCR. Test plates used in the transcriptional analysis were conducted on a mixture of equal volumes of two M. hyopneumoniae strain 232 RNA samples (stock 1) diluted to 4 ng μl⁻¹. A common non-inhibitory dilution series for the five or six target sequences within an IG region was determined separately for each region.

The qRT-PCR assays were conducted in a single 96-well plate on each of the two RNA samples in the non-inhibitory dilution range as determined in the test plates. The region ig-072 assay consisted of six primer sets and four dilutions per primer set. Each well condition was run in duplicate. Assays for regions ig-106 and ig-282 consisted of five primer sets and eight dilutions per primer set. Duplicate wells were run for primer set 3 in region ig-106 and for primer set 4 in region ig-282. Assays for regions ig-682 and ig-684 consisted of six primer sets and eight dilutions per primer set. No duplicate wells were run for these two regions.

Quantification cycle (Cq) values were obtained from the MxPro 4.1 software’s default amplification-based threshold method. This method specifies the threshold as the mid-point of the exponential amplification phase across all amplification curves on the plate. For duplicate samples on an assay plate (technical replicates), the Cq values were averaged for subsequent analysis.
### Table 1. Primers used for qRT-PCR analysis

<table>
<thead>
<tr>
<th>Region</th>
<th>Primer set</th>
<th>Forward primer</th>
<th>Sequence</th>
<th>Reverse primer</th>
<th>Sequence</th>
<th>Product size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ig-070/071</td>
<td>1</td>
<td>RT_uvrC.L</td>
<td>TGCTCCTTTTGCAATCAGTTCC</td>
<td>RT_uvrC.R</td>
<td>AAAATTGATGCTGCAAACC</td>
<td>121</td>
</tr>
<tr>
<td>ig-070/070</td>
<td>2</td>
<td>RT_uvrC-ig070.L</td>
<td>TTTGGACGCACTCAATTTT</td>
<td>RT_uvrC-ig070.R</td>
<td>TCACCTCAATTTATTTTAATCGAG</td>
<td>150</td>
</tr>
<tr>
<td>ig-070/071</td>
<td>3</td>
<td>RT_mhp071.L</td>
<td>TCTGTCCTTTTGCATACTGATTC</td>
<td>RT_mhp071.R</td>
<td>CGTAAATCTTCTGGCAAGC</td>
<td>142</td>
</tr>
<tr>
<td>ig-070/071</td>
<td>4</td>
<td>RT_mhp071-ig071.L</td>
<td>AAAAGTGGAATGATCTGAATACAATC</td>
<td>RT_mhp071-ig071.R</td>
<td>TGGCCAGATTTATGAAAATAGTCA</td>
<td>126</td>
</tr>
<tr>
<td>ig-070/071</td>
<td>5</td>
<td>RT_ig071-dnaK.L</td>
<td>TTTTCTAAAATTTTGGAATACCTTTT</td>
<td>RT_ig071-dnaK.R</td>
<td>CCTAAATTGATTTCTTTTGCAATT</td>
<td>147</td>
</tr>
<tr>
<td>ig-070/071</td>
<td>6</td>
<td>RT_dnaK.L</td>
<td>CTTGCTGCTTTTGAATATCC</td>
<td>RT_dnaK.R</td>
<td>CTCGCAATTTCTTCTTGG</td>
<td>82</td>
</tr>
<tr>
<td>ig-072*</td>
<td>1</td>
<td>RT_mhp072.F</td>
<td>ACAGTTGCTGCGAACAGGAAATTAAACATC</td>
<td>RT_mhp072.R</td>
<td>TGGGACGCTGCTTCAATTTTGTC</td>
<td>183</td>
</tr>
<tr>
<td>ig-072*</td>
<td>2</td>
<td>RT_mhp072.ig1.F</td>
<td>AAGAAATCCAGGAAGCTACCAATC</td>
<td>RT_mhp072.ig1.R</td>
<td>CTAACCAAGTTTGAATGTCG</td>
<td>129</td>
</tr>
<tr>
<td>ig-072*</td>
<td>3</td>
<td>RT_mhp072.ig2.F</td>
<td>AAAAGTGGAATGATCTGAATACAATC</td>
<td>RT_mhp072.ig2.R</td>
<td>AAGCCACTAGCACCAACCAACAA</td>
<td>112</td>
</tr>
<tr>
<td>ig-072*</td>
<td>4</td>
<td>RT_mhp072.ig3.F</td>
<td>TTTGGTCCAGGATGGTTTTATC</td>
<td>RT_mhp072.ig3.R</td>
<td>CGGCTGAAATCTTGAATTGTCACA</td>
<td>157</td>
</tr>
<tr>
<td>ig-072*</td>
<td>5</td>
<td>RT_mhp072.ig4.F</td>
<td>GTAACATCTGAAGGTTACGCACTGAC</td>
<td>RT_mhp072.ig4.R</td>
<td>TGGACCGCAACAAATTCCGGGAA</td>
<td>116</td>
</tr>
<tr>
<td>ig-072*</td>
<td>6</td>
<td>RT_mhp073.F</td>
<td>TTCTCCGCAATTTTGGAATACCTTAC</td>
<td>RT_mhp073.R</td>
<td>AACACATCTGCTGATGACGTCG</td>
<td>166</td>
</tr>
<tr>
<td>ig-106</td>
<td>1</td>
<td>RT_mhp106.F</td>
<td>TGAGTTGCAAGCAATTGATTCATG</td>
<td>RT_mhp106.R</td>
<td>CCAATATCTTCTTTAATTTG</td>
<td>200</td>
</tr>
<tr>
<td>ig-106</td>
<td>2</td>
<td>RT_mhp106.ig2.F</td>
<td>TGTCTCAGTTCAGTATGTTTCAAT</td>
<td>RT_mhp106.ig2.R</td>
<td>GGAATTATATAGTTATTCTTCAATTT</td>
<td>123</td>
</tr>
<tr>
<td>ig-106</td>
<td>3</td>
<td>RT_mhp106.ig3.F</td>
<td>ACAATTAATACCTTAAATTCCTT</td>
<td>RT_mhp106.ig3.R</td>
<td>AGAAGCACTAATTTCAAGA</td>
<td>82</td>
</tr>
<tr>
<td>ig-106</td>
<td>4</td>
<td>RT_mhp106.ig4.F</td>
<td>AATTGGAATGATTTTGGAATACCTT</td>
<td>RT_mhp106.ig4.R</td>
<td>CCAACGGGAAAATTTGACG</td>
<td>170</td>
</tr>
<tr>
<td>ig-106</td>
<td>5</td>
<td>RT_mhp107.F</td>
<td>GCAGGCTAATTGATTTCTTAC</td>
<td>RT_mhp107.R</td>
<td>CCGGCGGGAACCAATTTATTTAT</td>
<td>139</td>
</tr>
<tr>
<td>ig-282*</td>
<td>1</td>
<td>RT_mhp282.F</td>
<td>CCATAGTGAAGCAGATGTTTACGAT</td>
<td>RT_mhp282.R</td>
<td>TGGGACGCTGCTTCAATTTTGTC</td>
<td>91</td>
</tr>
<tr>
<td>ig-282*</td>
<td>2</td>
<td>RT_mhp282.ig2.F</td>
<td>GATGGATGTAAGTGAGTTTGGGCA</td>
<td>RT_mhp282.ig2.R</td>
<td>TTTCCGCAATTTGACG</td>
<td>88</td>
</tr>
<tr>
<td>ig-282*</td>
<td>3</td>
<td>RT_mhp282.ig3.F</td>
<td>CTCTGTCCTTTTGAATATCC</td>
<td>RT_mhp282.ig3.R</td>
<td>AATATTGATCTTCTAATTTTCAACAA</td>
<td>115</td>
</tr>
<tr>
<td>ig-282</td>
<td>4</td>
<td>RT_mhp282.ig4.F</td>
<td>GACATTATTTGATGCTT</td>
<td>RT_mhp282.ig4.R</td>
<td>GCCGTTGATCTAATTTCAAGT</td>
<td>153</td>
</tr>
<tr>
<td>ig-282</td>
<td>5</td>
<td>RT_mhp283.F</td>
<td>AATTGGAATGATTTTGGAATACCTTAC</td>
<td>RT_mhp283.R</td>
<td>TTGCTGCCTTCTTTGCAATTTAT</td>
<td>90</td>
</tr>
<tr>
<td>ig-682</td>
<td>1</td>
<td>RT_mhp682.F</td>
<td>AGTGACTCAAGTTGATGTTTGAAGA</td>
<td>RT_mhp682.R</td>
<td>ATAAAGCTCGCTGCTAGG</td>
<td>199</td>
</tr>
<tr>
<td>ig-682*</td>
<td>2</td>
<td>RT_mhp682.ig1.F</td>
<td>AACTTGAACATCAAAGGGAATACGA</td>
<td>RT_mhp682.ig1.R</td>
<td>GTTTAATGCTGCACTGAAACT</td>
<td>142</td>
</tr>
<tr>
<td>ig-682*</td>
<td>3</td>
<td>RT_mhp682.ig2.F</td>
<td>TCTGGTCCTTTTGAATATCC</td>
<td>RT_mhp682.ig2.R</td>
<td>GGGCTGAAATGTGTTGATT</td>
<td>179</td>
</tr>
<tr>
<td>ig-682</td>
<td>4</td>
<td>RT_mhp682.ig3.F</td>
<td>CAGCGCACAATCAATCATAA</td>
<td>RT_mhp682.ig3.R</td>
<td>CATATATCGTAAGGTCAG</td>
<td>118</td>
</tr>
<tr>
<td>ig-682*</td>
<td>5</td>
<td>RT_mhp682.ig4.F</td>
<td>TGACGCGGATCTAGC</td>
<td>RT_mhp682.ig4.R</td>
<td>GGGCATTATGAGGGAATAAA</td>
<td>170</td>
</tr>
<tr>
<td>ig-682*</td>
<td>6</td>
<td>RT_mhp683.F</td>
<td>TGGGCTGCTTTGGAATATC</td>
<td>RT_mhp683.R</td>
<td>ATGATAGTCTGCTTGACGCAGG</td>
<td>136</td>
</tr>
<tr>
<td>ig-684*</td>
<td>1</td>
<td>RT_mhp684.F</td>
<td>AACTTGAACATCAAAGGGAATACGA</td>
<td>RT_mhp684.R</td>
<td>ATGATAGTCTGCTTGACGCAGG</td>
<td>98</td>
</tr>
<tr>
<td>ig-684</td>
<td>2</td>
<td>RT_mhp684.ig1.F</td>
<td>ACAATTTGAATCTGATGTTTGAAGA</td>
<td>RT_mhp684.ig1.R</td>
<td>TGGGCTGAAATGTGTTGATT</td>
<td>116</td>
</tr>
<tr>
<td>ig-684*</td>
<td>3</td>
<td>RT_mhp684.ig2.F</td>
<td>AACTTGAACATCAAAGGGAATACGA</td>
<td>RT_mhp684.ig2.R</td>
<td>GTTTAATGCTGCACTGAAACT</td>
<td>135</td>
</tr>
<tr>
<td>ig-684</td>
<td>4</td>
<td>RT_mhp684.ig3.F</td>
<td>GTTGGTCTTAACCTCCTATTAT</td>
<td>RT_mhp684.ig3.R</td>
<td>ATATCGTTAAGTATTTGAGGTTGATAT</td>
<td>104</td>
</tr>
<tr>
<td>ig-684</td>
<td>5</td>
<td>RT_mhp684.ig4.F</td>
<td>TATAATGAGGTAATGGAATACGA</td>
<td>RT_mhp684.ig4.R</td>
<td>GGTCGTAAGAATCCGCACG</td>
<td>200</td>
</tr>
<tr>
<td>Reference</td>
<td>1</td>
<td>RT_mhp381.L2</td>
<td>TGACATATAATCTGGCAGAATTC</td>
<td>RT_mhp381.R</td>
<td>ACTGGTGAATATCCTGGAAGG</td>
<td>172</td>
</tr>
</tbody>
</table>

*Indicates primers were confirmed with DNA template in qRT-PCR.
qRT-PCR relative quantity index. A simple linear model of 
\( C_q = mx + b \), where \( C_q \) is the quantification cycle and \( x = \log_2(z) \),
where \( z \) is the dilution factor, was fitted to the data of each primer set
in an assay. Note that the dilution factor \( z \) (and \( x \)) is inversely
proportional to the concentration of target RNA in the sample; as the
sample dilution increases, the concentration of the target sequence RNA decreases. A standard curve is constructed by estimating \( m \) and \( b \)
from the data of a given dilution series. Extrapolating this standard curve to \( C_q = 0 \), the cycle preceding PCR where a sample would be at
the detection level, the model becomes \( 0 = mx + b \). Solving for \( x \) results
in \( x = -b/m \). This quantity is a relative measure of the initial amount
of target sequence RNA in the sample, but since \( x \) is inversely related
to relative quantity, \( -b/m \) is also inversely related to relative quantity.
Thus, a lower \( -b/m \) corresponds to more target RNA in the sample.
An index that would be more intuitive is the negative of
\( -b/m \), namely \( b/m \). This index is directly related to relative quantity. That is,
the higher the \( b/m \), the more target sequence RNA there is in the
sample. A derivation of this relative quantity may also be seen as follows:
\( b/m = -x = -\log_2(z) = \log_2(z^{-1}) = \log_2(Q) \), where \( Q \) is the
relative quantity measure of the target RNA in a sample. An example
calculation of the relative quantity index is available with the online
version of this paper (Supplementary Methods: example calculation
of the relative quantity index).

qRT-PCR DNA assay analysis. A single plate and sample
DNA was used in the assay. Separate linear models were fitted to each primer set
to obtain estimates of \( m \) and \( b \). These estimates were used to estimate
the relative quantity index \( (b/m) \), and the delta method (Casella &
Berger, 2002) was employed to obtain an estimate of the \( \text{SE} \) of these
quantity indices. All pairwise comparisons of the six primer set
indices were done with a Bonferroni correction for multiple testing.

qRT-PCR transcript analysis. For the IG region assays with two
RNA samples, the simple linear model above was extended to a linear
model with a separate mean for each RNA sample: 
\( C_q = mx - b_i \), where \( b_i \) is the intercept for RNA sample \( i \). Thus, the model
assumes that the two RNA samples have the same slope \( (m) \) but different
intercepts. For primer sets that had technical replications, the linear
model was fitted to the averaged \( C_q \) values. Separate linear models
were fitted to each primer set to obtain estimates of \( m \) and \( b_i \) for each
RNA sample. The relative quantity index \( (b/m) \) was estimated from
these results and the delta method was employed to obtain an
estimate of the \( \text{SE} \) of these quantity indices. All pairwise comparisons
of the six primer set indices were done with a Bonferroni correction
for multiple testing. To identify stem-loop structures in the IG-072,
IG-106, IG-282, IG-682 and IG-684 IG regions, SoftBerry FindTerm
software was used.

Heat-shock experimental design and data analysis. The
upstream region (ig-070/071) and downstream region (ig-072) of
dnaK served as a positive control since previous studies have shown that
dnaK is upregulated during heat shock (Madsen et al., 2006b).
Six RNA samples, three control and three heat shock, were used in the
assays. Test plates for the RNA samples and primer sets for both
regions on either side of dnaK were conducted on equal volumes of
three heat-shock RNA samples and three control RNA samples
diluted to 4 ng μl⁻¹ (stock I). Once the non-inhibitory dilution range
for each region had been determined across the six target sequences in
each region, the qRT-PCR assay was run in duplicate. Half of an assay
plate consisted of six dilutions of stock I (individual wells for the
target sequences and duplicates for the reference gene) in order to
construct the standard curves. The other half of the assay plate
consisted of each RNA sample combined with the six primer sets run
individually and a seventh primer set for the reference gene (mhp381)
run in duplicate. In the downstream region (ig-072), the three heat-
shock RNA samples and three control RNA samples were diluted 1:3
for an in-well dilution of 1:9; and for the region upstream of dnaK
(ig-070/071), these samples were diluted 1:5 for an in-well dilution of
1:15. These sample dilutions corresponded near to the centre of the
stock I dilution series as determined from the test plates for each
region.

The standard curves for each probe set were calculated from the stock
I dilution series by fitting the linear model \( C_q = mx + b \) to obtain estimates of \( m \) and \( b \) for each primer set per plate. The observed \( C_q \)
values for the RNA samples were then normalized \( (C_q') \) by the
reference gene (see Supplementary Methods for an example
calculation). After normalization, the relative quantity of a primer
set was then calculated as \( C_q'/m \). This relative quantity is a linear shift
of the relative quantity index \( b/m \). The derivation is easily seen by
recalling that the linear model is \( C_q = mx + b \), so \( C_q/m = (mx + b)/m \)
\( = x + b/m \). Thus, for each target sequence there are two sets of relative
indices, one set for the control samples and the other set for
the heat-shock samples. These indices are then averaged within
treatment groups and the \( \text{SEM} \) for each target sequence/treatment
combination is calculated. Fold change was calculated as described by
the ISU equation of Gallup & Ackermann (2006).

RESULTS

Microarray design
An analysis of the IG regions of the annotated M.
hyopneumoniae strain 232 genome (Minion et al., 2004)
defined 215 IG regions of 125 bp or more and 128 IG
regions of 50–124 bp for a total of 343 IG regions (see
Supplementary Fig. S1). The remaining 230 IG regions of
less than 50 bp were not analysed since previous studies
have shown that regions smaller than 50 bp could serve to
separate ORFs in an operon from which a polycistronic mRNA
is produced (Adams et al., 2005). A total of 158 PCR-
generated probes for 112 of the 215 IG regions larger than
125 bp were designed. Some of the larger regions had more
than one probe designed to achieve at least 50% coverage.
An additional 517 oligonucleotides for 247 IG regions
were designed covering 128 IG regions of 50–124 bp and
119 IG regions of 125 bp or more for which PCR primers
could not be designed due to the high A+T content of
the IG regions (81.3 mol%) or a failure of PCRs (after
repeated primer designs) for unknown reasons. A total of
16 IG regions had both PCR and oligonucleotide probes
constructed.

Microarray IG region probe detection
Since mycoplasmal mRNAs lack polyadenylated sequences,
we sought the most efficient primer set for producing
cDNAs from RNA preparations. Our choices were a full
random primer set of hexamers and a subset of 129
hexamer primers shown to bind to protein-encoding M.
hyopneumoniae ORFs used in previous microarray studies,
and listed in Supplementary Table S3 of Madsen et al.
(2006b). We compared the two using the same RNA
preparation at identical concentrations in reverse tran-
scriptase reactions to generate cDNA for subsequent
labelling reactions. Our hypothesis was that the random
primers would be more efficient in producing cDNAs from

http://mic.sgmjournals.org
IG RNAs than the subset of primers selected for ORF sequences and consequently produce stronger signals on the IG arrays. One caveat that might confound that hypothesis is the fact that we would be labelling rRNAs as well. A visual inspection of both arrays following hybridization with the two labelled cDNA preparations showed no appreciable difference in probe fluorescence between the two types of primers used to generate cDNA (data not shown). Therefore, the data from both primer sets were used in the analysis of the array features.

Applying the detection criterion (more than 6 SEM above the mean intensity of spot buffer-only probes in both channels) to the scanned arrays resulted in most of the IG regions declared as being transcribed. For the PCR probe array 94 out of 112 (83.9%) IG regions had positive signals, and for the oligonucleotide array 239 out of 247 (96.8%) IG regions had positive signals. Not all probes within an IG region had a positive signal, but for a region to be declared positive, at least one probe within a region was positive. Since 16 IG regions had both kinds of probes, of which 12 detected positive signals, this resulted in a total of 321 (94 PCR + 239 oligonucleotide – 12 double counted) out of the 343 IG regions (93.6%) with positive signals across both array types. Interestingly, of the 73 IG regions with outwardly divergent transcription, 68 (93.2%) had positive signals. Those IG regions with inward transcription from both ORFs had 98% positive signals (52/53), and the IG regions with at least one inward transcript had 92.6% (201/217) positive signals (see Supplementary Table S1). Analysis of five IG regions for stem–loop structures identified potential stem–loops in ig-072 at the end of mhp072 and in ig-682 at the end of mhp683 (Fig. 1b, e). However, these putative structures lacked sufficient G+C content, suggesting that transcription termination is ineffective in the typical rho-independent fashion (Farnham & Platt, 1981).

RT-PCR DNA assay

*M. hyopneumoniae* chromosomal DNA template was used as a negative control to assess the relative quantity index and accuracy of the qRT-PCR. In this assay, no significant differences between the six target sequences in region ig-072 would be expected since the relative concentration of template should be the same across all areas of the IG and flanking regions. As expected, no differences were observed across all pairwise comparisons of six probes (15 tests of significance) at P<0.01 (Fig. 1a).

RT-PCR IG regions assay

Thirty-five IG regions greater than 500 bases in length were identified in our analysis, and from these, five were selected for qRT-PCR analysis based on the direction of transcription of the flanking ORFs. Representative regions of the four possible read direction combinations were selected with the following criteria in mind: (1) one IG region had both flanking ORFs on the forward strand (ig-072); (2) one IG region had both flanking ORFs on the reverse strand (ig-282); (3) one region had both flanking ORFs transcribed into the IG region (ig-682); and (4) two regions had both flanking ORFs divergently transcribed away from the IG region (ig-106, ig-684). A sixth region of less than 500 bp comprising the upstream region of dnaK was examined as a positive control; this region is known to undergo significant upregulation during heat shock (Madsen et al., 2006b).

The five IG regions were assayed using two biologically distinct RNA samples. The pattern of relative transcription was nearly identical for both RNA samples, as can be seen in Fig. 1(b–f). Statistically significant differential levels of transcription were detected in four out of five regions. The most common difference was a drop-off in RNA levels at ORF/intergenic transitions. In only one region, ig-682, did we fail to detect a signal by qRT-PCR at one of the six target sequences. To confirm the primers and reaction conditions for this target sequence, for other selected primer sets (Table 1), and for the entire ig-072 region, *M. hyopneumoniae* chromosomal DNA was used as a template in the qRT-PCR, which gave positive reactions for every primer set tested (data not shown). In Fig. 1(a) we show an example of the chromosomal DNA control reactions for region ig-072.

Analysis of region ig-072 (Fig. 1b) shows that for both biological samples, the qRT-PCRs that bridge the transition from ORF to IG region (primer sets 2 and 5) have significantly lower amounts of transcripts (P<0.02) than the adjacent ORF regions (primer sets 1 and 6) although some signal was detected. There was no significant difference between the IG probes (primer sets 3 and 4) and the two ORF probes (primer sets 1 and 6). Following the decrease in transcript levels at the end of mhp072 (primer set 2), there was an increase in transcript levels across the IG region (primer sets 3 and 4) followed by another decrease at the ig-072–mhp073 junction (primer set 5). This increase was statistically significant (P<0.01) between primer sets 2 versus 3 and 4 for one of the two RNA samples and the decrease was significant for both RNA samples (primer sets 3 and 4 versus 5).

Region ig-106 (Fig. 1c), which contains sequences between divergently transcribed (outward) genes mhp106 and mhp107, shows a different pattern of transcription. The IG region (primer sets 2 and 3) demonstrates a statistically indistinguishable level of transcription (P=0.1) relative to mhp107 (primer set 5) for one of the two RNA samples. Transcription in mhp106 (primer set 1), however, is significantly higher (P<0.01) than that in the IG region (primer sets 2 and 3).

Region ig-282 (Fig. 1d), which has both ORFs transcribed in the same direction on the reverse strand relative to DNA replication, also demonstrated IG transcripts. At the end of mhp283 (primer set 4), transcription decreased significantly (P<0.0001) from levels within mhp283 (primer set 5) and then increased within the IG region (primer sets 2 and 3).
Intergenic transcription in *M. hyopneumoniae*

Fig. 1. qRT-PCR analysis of IG regions of *M. hyopneumoniae*. Shown are the relative quantity (b/m) estimates of each qRT-PCR plotted vertically over the corresponding product location in the *M. hyopneumoniae* genome sequence given in base pairs. The data represent the relative quantity index estimate ± 2 SE of the estimate. The IG designation is given at the top of each panel. The circles and squares depict results for the two biologically distinct samples. Each curve represents an independent RNA sample except for (a), where genomic DNA was used as a template for a positive control reaction. The letters indicate statistically significant differences. Points with the same letter are not significantly different (*P* < 0.01) for that independent RNA sample. The ORF designation and gene name are given for each flanking gene (H, hypothetical), and the arrows indicate the direction of transcription. The open bars indicate the location of PCR products used as probes in the microarrays. The filled bars indicate the qRT-PCR target sequences, and the numbers indicate the qRT-PCR primer sets described in Table 1 in that region. ND indicates that the probe was not detected in the assay. Ω, putative stem–loop structure as predicted by SoftBerry FindTerm software.

3, *P* < 0.0001). Transcription within mhp282 (primer set 1) was higher (*P* < 0.0001) than in the IG region (primer sets 2 and 3).

For region ig-682 (Fig. 1e), where flanking ORFs transcribe into the IG region, transcript levels between mhp682 (primer set 1) and the bridge region (primer set 2) were the
same followed by a decrease \( (P<0.01) \) in the IG region (primer set 3). No signal was detected with the ig-682/mhp683 bridging region (primer set 5) even though the DNA template control was positive (data not shown). Transcription in mhp683 (primer set 6) was higher than that of all other areas within this region \( (P<0.01) \).

The primer sets for region ig-684 (Fig. 1f) did not perform well, as indicated by the higher level of variation. This region is of the same type as ig-106, with both flanking ORFs being transcribed away from the IG region, but the region has a higher A+T content (83.3 mol%), significantly contributing to difficulty in primer design. The trend in transcript levels of ig-684 is similar to ig-106, where the IG region had detectable but lower levels of transcription relative to either of the ORFs.

**RT-PCR analysis of the dnaK up- and downstream regions following heat shock**

The region surrounding dnaK was studied as a positive control because both dnaK and dnaJ have been shown to be highly upregulated during heat shock (Chang et al., 2008; Madsen et al., 2006b). For this study, IG regions upstream (ig-070/071) and downstream (ig-072) of dnaK were assayed separately. The results can be seen in Fig. 2.

The top graph in Fig. 2 shows the log₂ fold change in transcription levels between control and heat-shock conditions. All target sequences in both regions indicate increased levels of mRNA under heat-shock conditions. This increase was statistically significant \( (P<0.05) \) for all primer sets except for primer set 3 in ig-070/071 \( (P=0.11) \). In region ig-070/071, the dnaK target sequence was the most upregulated: 13.38-fold higher amount of transcript than under normal growth conditions. All other target sequences in this region were upregulated about 2-fold. In region ig-072, the mhp072 bridge target sequence had the highest level of upregulation (20.95-fold higher transcript levels), with dnaK as the next highest level of transcript increase (7.26-fold). The rest of region ig-072 was upregulated 3–4-fold, except for the other bridge area with dnaJ, which was upregulated about 5-fold.

The lower graph depicts the relative index \( (Cq/m) \) averaged across the three biological replicates within each treatment for both regions. Primer sets 1, 2 and 5 in region ig-070/071 have lower levels than primer sets 3, 4 and 6 \( (P<0.01) \) for both treatment groups. The results for region

![Fig. 2. Transcriptional analysis of IG regions flanking dnaK following heat shock. The map shows the corresponding region in the genome of M. hyopneumoniae given in base pairs. The ORF designation and gene name is given for each gene, and the arrows indicate the direction of transcription. mhp071 is a hypothetical gene. Open bars indicate the location of probes (PCR products or oligonucleotides) used in the microarrays. Filled bars indicate the qRT-PCR target sequences, and the numbers indicate the qRT-PCR primer sets described in Table 1 in that region. The upper panel gives the fold change (mean ± 2 SEM) of triplicate samples. The asterisks indicate significant differences between the control and heat-shocked samples \( (*, P<0.05; **, P<0.01) \). The lower panel gives the relative quantity \( (Cq/m) \) estimates of each qRT-PCR at the two temperatures. Control samples are represented by circles and heat-shocked samples by squares. Data represent the mean ±2 SEM of triplicate samples (note: most SEM bars are not visible due to being smaller than the plotted point). The letters indicate significant differences. Points with the same letter are not significantly different \( (P>0.01) \) for that treatment.

S. W. Gardner and F. C. Minion

Microbiology 156
ig-072 are very similar to the results from the qRT-PCR region assay described above. For both treatment groups, the bridge areas have an appreciably lower level of transcription, and the IG region has levels of transcription closer to the dnaJ levels.

**DISCUSSION**

This study represents a global assessment of transcription within IG regions of the *M. hyopneumoniae* genome. We used both microarrays and qRT-PCR to assess transcription. The decision was made to limit our analysis to IG regions of 50 bp or more using a combination of PCR products and oligonucleotide as probes. Of the IG regions in this set, 35 were large (>500 bp), requiring multiple PCR products or oligonucleotides for at least 50 % coverage. Coverage was not 100 % because of the inability to design unique primers or oligonucleotides in regions of high A+T content.

We considered the possibility in our primer design that cross-hybridization between ORF transcripts and IG region probes could account for some of the signals we obtained, and without more extensive analysis could not totally discount that possibility. Our detailed analysis of five IG regions by qRT-PCR, however, showed no evidence of cross-hybridization effects. All melt curves from qRT-PCR assays indicated that a single product was generated per target sequence.

Our first study with the arrays was to determine the optimal conditions for generating cDNA for labelling reactions using two different primer sets. The hexamer subset was chosen based on their lack of binding to *M. hyopneumoniae* rRNAs and binding to annotated ORFs in the genome (Madsen et al., 2006b). It was initially thought that this primer set might not label RNAs from IG regions and consequently give false negative signals. The second primer set was a commercial random hexamer primer preparation that would label rRNAs representing about 90 % of the total RNAs, diluting the signals on the arrays while also labelling IG region RNAs. Unexpectedly, both primer sets gave similar signals across both IG arrays (data not shown). Thus, we used both primer sets in our analysis of the IG regions transcripts. A posteriori, it seems that equal labelling efficiency with the two primer sets may be due to runoff transcription into IG regions.

Our results indicated that a large proportion of the IG regions, 321 out 343 (93.6 %), were transcribed. This was unexpected and indicated that *M. hyopneumoniae* does not strictly control transcription termination as previously suggested (Washio et al., 1998). These studies are supported by recent studies with *M. pneumoniae* in which IG transcripts on a global scale were also identified by tiling microarrays (Güell et al., 2009) and on a limited scale with *Mycoplasma genitalium* and *M. pneumoniae* (Benders et al., 2005). From a physiological standpoint, the lack of transcription termination seems energetically wasteful, suggesting that mycoplasmas may have important roles for these transcripts, such as the generation of small regulatory peptides or RNAs that have yet to be defined in these organisms. None of these IG regions had definable promoter-like sequences using current algorithms (data not shown), although promoters in mycoplasmas have not been fully defined by mutational analysis and could vary between species and not be recognized by current software. Stem–loop structures that may define rho-independent transcriptional termination sites are thought not to be important in mycoplasmas (Washio et al., 1998). Our analysis of the five IG regions and adjacent ORFs examined showed evidence of weak stem–loop structures in two regions, ig-072 and ig-682 (data not shown).

The five regions for qRT-PCR analysis were chosen based on size (>500 bp), direction of flanking ORF transcription (similar and divergent transcription) (Fig. 1) and in the case of dnaK its known transcriptional control mechanism (Chang et al., 2008). The chosen IG regions also represented both DNA replication forks to rule out any replication effects on transcription. In most bacteria, the direction of transcription and DNA replication coincide in most genes. Even though this is not true for *M. hyopneumoniae* (Minion et al., 2004), we wanted to eliminate this possibility by choosing IG regions from both sides of the chromosome. Regions ig-072, ig-106 and ig-282 were associated with the right clockwise replication fork, and regions ig-682 and ig-684 with the left counter-clockwise replication fork. Also, mhp107 and mhp684 are P97 paralogues (Adams et al., 2005; Minion et al., 2004), and mhp683 is a P102 parologue. These proteins, along with other members of these families, are thought to play roles in adherence to swine epithelium, enhance colonization and thereby contribute to virulence (Adams et al., 2005; Burnett et al., 2006; Wilton et al., 2004, 2009), thus increasing our interest in them for defined analysis.

Before analysing the data, we considered two possibilities: an abrupt termination of transcription at the end of the flanking genes with the loss of bridging qRT-PCR products (mechanism 1) and a general decline in transcript levels indicating that RNA polymerase was slowly falling off the template in the IG region (mechanism 2). For regions with outward divergent transcription, we thought transcription would not occur. Our data indicate that both mechanisms occur along with a third mechanism: reinitiation of transcription within the IG region (mechanism 3).

Analysis of five IG regions showed a general trend of lower transcription in the IG region relative to the flanking ORFs (Fig. 1). It is important to note that none of these regions was potentially involved in an operon producing a polycistronic message. This implies that there is no obvious reason for continued transcription across the region. However, it is clear from these studies that transcription initiated (mechanism 3) within IG regions ig-072, ig-106 and ig-282 since downstream of the flanking ORF the level of transcripts increased significantly (Fig. 1b–d). This is
independent of the direction of DNA replication; the genes flanking ig-072 and ig-282 are transcribed in the same direction but on opposite strands. There is a considerable reduction in transcript levels at the end of the first gene, a general rise in transcript levels in the IG region and increased transcription in the downstream gene. This implies a third mechanism, one in which transcription never truly ceases between genes and reinitiation of transcription occurs within the IG region before encountering the promoter of the second flanking ORF.

Region ig-682 differed, however. This is a region where flanking ORFs are inwardly divergently transcribed (Fig. 1e). From mhp682, transcription continued into the IG region, and it appeared that the RNA polymerase gradually fell off the template as it proceeded into the IG region (mechanism 2). This is the only region studied that showed the same transcript levels within both the gene (primer set 1) and the bridging region (primer set 2). For mhp683, however, there was an abrupt termination of transcription at the end of the gene (mechanism 1). This was the only gene in which transcription was completely terminated; most of the other genes showed decreased transcription levels in bridging regions but significantly higher levels of transcription within the IG region (Figs 1 and 2) or IG region transcripts continuing into the flanking ORF. Further analysis of the IG region adjacent to mhp683, however, failed to reveal a sequence suitable for rho-independent transcription termination as defined in other genetic systems; although a stem–loop structure could be identified just downstream of the left primer binding site for primer set 5 in ig-682 (Fig. 1e), the region was A+T-rich, lacking the characteristic G+C content of a transcription terminator. If this structure serves the purpose of a termination signal in \(M. \text{hyopneumoniae}\), it is unusual. Analysis of IG regions ig-072 at the end of mhp072 and ig-282 at the end of mhp283 revealed smaller stem–loop structures also lacking G+C content, suggesting that transcription termination is ineffective in the typical rho-independent fashion.

In further analysis of the heat-shock gene \(dnak\) and its flanking regions by qRT-PCR, we observed a general increase in transcription across the entire region during heat shock (Fig. 2). This region was included because of previous studies demonstrating increased transcription of \(dnak\) during heat shock (Chang et al., 2008; Madsen et al., 2006b), and positive control of transcription changes was thought to be informative. Interestingly, there was a significant difference between control and heat-shock samples with every primer set tested except one. Because transcript profiles between control and heat-shocked samples were identical in the IG regions ig-070/071 and ig-072, this suggests that increased temperatures have an overall positive effect on transcription, at least in this genomic region. As expected, there was a much larger increase in mhp072 (\(dnak\)) transcripts (Chang et al., 2008; Madsen et al., 2006b), and previous studies showed that this gene is controlled by the heat-shock regulator \(HrcA\) (Chang et al., 2008). Also as expected, there was a sharp decrease in transcription at the end of \(dnak\), but some transcript remained, suggesting that termination was not complete. We were also able to show that transcript levels vary between genes: mhp070 (\(uvrC\)) had significantly (\(P<0.01\)) lower transcript levels than mhp071 and mhp072 (\(dnak\)) under both growth conditions.

In summary, these results indicate that IG regions within the \(M. \text{hyopneumoniae}\) genome are transcriptionally active. For many regions, transcription continues from flanking genes with a gradual reduction until the next promoter is encountered. For some regions, there is an internal initiation of transcription from ill-defined sequences. Most surprising were the observations that almost every IG region could be detected on the arrays, suggesting that IG regions in mycoplasmas may hold surprises with functions yet to be determined.

ACKNOWLEDGEMENTS

This study was funded partially by a grant from the Iowa Livestock Health Advisory Council from the College of Veterinary Medicine, Iowa State University, and from a US Department of Agriculture Formula Fund to F.C.M. We thank Jack Gallup and Michael Carruthers for technical assistance with the qRT-PCR assay, and Melissa Madsen for help with designing the probes, microarray assays and reviewing the manuscript. We also thank Stephen Windom and Sunny Gulati for their assistance in writing computer programs to process the data.

REFERENCES


Burnett, T. A., Dinkla, K., Rohde, M., Chhatwal, G. S., Upphoff, C., Srivastava, M., Cordwell, S. J., Geary, S., Liao, X. & other authors (2006). \(P159\) is a proteolytically processed, surface adhesin of \(M. \text{hyopneumoniae}\), defined domains of \(P159\) bind heparin and promote adherence to eukaryotic cells. Mol Microbiol 60, 669–686.


Edited by: C. Citti