Investigating the relationship between mRNA degradation rates and secondary structure in mycobacteria

A Major Qualifying Project Report

Submitted to

WORCESTER POLYTECHNIC INSTITUTE,

In Partial Fulfillment of the Requirements for the

Degree of Bachelor of Science In

Biology & Biotechnology and Bioinformatics & Computational Biology

Authored by:
Jessica Kelly
May 5, 2021

Approved by:
Scarlet S. Shell, PhD
Biology and Biotechnology
Bioinformatics and Computational Biology

Abstract

Mycobacterium tuberculosis (MTB) infections cause tuberculosis, one of the top ten causes of death worldwide. To survive stress conditions within the host, MTB regulates its transcriptome, which includes regulating mRNA degradation rates. In *E. coli,* it has been shown that secondary structures in 5' Untranslated Regions (5' UTRs) decrease the efficiency of mRNA degradation by hindering RppH and RNase E activity. In mycobacteria, the mechanisms regulating mRNA degradation are poorly understood, despite the presence of an RNase E homolog. In MTB, the mRNA coding for the virulence factor EsxB exhibits increased half-life following a cleavage event, relative to its parent transcript. We hypothesized that the increased stability is caused by formation of two terminal hairpin loops in the *esxB* 5' UTR post-cleavage. Surprisingly, deletion of the *esxB* 5' UTR loop structures did not increase the transcript degradation rate. Secondary structures in 5' UTRs were also predicted computationally transcriptome-wide, and base pairing probabilities were calculated to identify structured regions. Similarly, no significant relationship between base pairing in 5' UTRs and mRNA half-life were revealed in MTB or *Mycobacterium smegmatis*. However, unpaired regions in the ribosome binding site and start codon had a significant positive correlation with mRNA stability and ribosome binding, suggesting that translation efficiency may play a substantial role in controlling mRNA stability.

Introduction

Tuberculosis (TB) is one of the top ten causes of death worldwide, having killed an estimated 1.4 million people in 2018 (World Health Organization, 2020). TB is difficult to treat, requiring an intensive 6-month multidrug regimen. An additional challenge to treating TB is the evolution of multi-drug resistant strains of *Mycobacterium tuberculosis* (MTB), the bacteria that causes TB, which are able to withstand the standard antibiotic treatment. Novel antibiotics and other therapies aimed to treat TB more effectively and efficiently are crucial. However, a better understanding of how MTB regulates gene expression in mechanisms of virulence and environmental stress response is necessary.

Transcript degradation is an important, although poorly understood factor involved in bacterial stress response mechanisms. mRNA degradation helps determine the amount of transcript available for translation, thus directly affecting protein expression levels. Research has shown that MTB globally stabilizes its transcriptome in response to environmental stress, suggesting regulation of mRNA stability is crucial in its ability to survive the host environment (Rustad *et al*., 2013). Decay of mRNA is mediated by a degradosome made up of RNases, first via cleavage by endoribonucleases, followed by rapid digestion by exoribonucleases (reviewed in Rauhut & Klug, 1999). In *E. coli,* the best studied bacterial system regarding RNA stability, the endoribonuclease RNase E provides the rate-limiting step in RNA degradation, preferentially binding to single-stranded RNA (Cohen *et al*., 1997). The MTB degradosome contains a homolog of RNase E, as well as PNPase and RNase J, an enzyme with both endonucleolytic and 5' to 3' exoribonuclease activity (Taverniti *et al*., 2019).

In *E. coli*, RNase E begins scanning at the 5' end of a transcript to catalyze downstream cleavage, dependent on the presence of a 5' monophosphorylated terminus (Celesnik *et al*., 2007). The enzyme RppH in *E. coli* is crucial in removal of a 5' pyrophosphate to create the

monophosphorylated terminus RNase E needs for efficient degradation (Deana *et al*., 2008). The ability of RNases to cleave mRNA is dependent on several factors that influence binding ability. This includes small RNAs, ribosome binding (Braun *et al*., 1998), 5' triphosphate groups (Bouvet & Belasco, 1992) and secondary structure in *E. coli*, ribosome binding (Condon 2003) and secondary structure (Condon 2003, Hubraeus *et al*., 2000) in *Bacillus subtilis*, and leader/leaderless status in *M. smegmatis* (Nguyen *et al*., 2020).

Our particular interest is investigating the relationship between transcript stability and secondary structure in 5' regions. Hairpin and stem loops near the 5' end of a transcript have been shown to impede mRNA degradation in several contexts. Literature reports deletion of the native 5' terminal stem-loop structure in the *ompA* mRNA, one of the most stable mRNAs in *E. coli*, decreases its half-life by a factor of three (Emory *et al*., 1992). The same study showed the addition of a stem loop to the 5' terminus doubles the mRNA half-life of *bla* transcript (Emory *et al*., 1992). A similar example in *B. subtilis* was observed by deleting the 5' terminal stem loop in the stable *aprE* mRNA to reduce its half-life by a factor of five (Humbraeus *et al*., 2000). Other examples of stability caused by 5' secondary structures have been identified in *Rhodobacter capsulatus* (Klug *et al*., 1990), *Acinetobacter baumannii* (Ching *et al*., 2017), and *Helicobacter pylori* (Amilon *et al*., 2015).

Secondary structures at the 5' terminus disrupt the ability of RNase E activity in one of two proposed ways: impeding its direct binding at the 5' terminus (Carrier *et al*., 1999, Condon 2003) or preventing RppH from creating the monophosphorylated 5' terminus to stimulate RNase E activity (Deana *et al*., 2008). In *E. coli*, the proximity of structures to the 5' end of transcripts has proven to be significant, as mutated transcripts with more than five free nucleotides at the 5' terminus were no longer stabilized by 5' stem loops in two studies (Emory *et al*., 1992, Bouvet & Belasco, 1992). In addition, RppH activity in *E. coli* was shown to be nine times greater when there was at least one nucleotide unpaired at the 5' terminus (Deana *et al*., 2008).

In addition to inherent structures that form upon transcription, there are also situations such as riboswitches, in which mRNA structure may be altered as a form of transcriptional or post-transcriptional regulation. Riboswitches typically use binding of small molecules or changes in temperature to cause changes in mRNA structure that allow or prevent binding of the ribosome (Nahvi *et al*., 2002). An example reported in *E. coli* shows that the *dinQ* mRNA, which is inherently unable to be translated due to secondary structure that blocks the Shine-Dalgarno (SD) site, undergoes conformational change. The change in secondary structure of *dinQ* is dependent on binding of an antisense RNA which disassembles a hairpin loop at the SD sequence and reforms the loop further upstream, both stabilizing the transcript and permitting translation (Kristiansen *et al.*, 2016).

There are reports in the literature suggesting that, although uncommon, mRNA processing may in some cases play a role in selective transcript stabilization to regulate protein expression levels (Obana *et al*., 2010, Sala *et al*., 2008). Within the MTB transcriptome, one mRNA processing site exists in a transcript encoding components of the virulence-associated ESX-1 secretion system (Shell *et al*., submitted manuscript). This system includes a 1:1 heterodimeric complex of secreted virulence factors EsxB and EsxA, which were shown to have roles in bacterial cell wall integrity (Garces *et al*., 2010) and bacterial translocation to the cytosol within macrophages (Ma

et al., 2015). Without proper formation of the ESX-1 secretion system, MTB has reduced pathogenicity (Pym *et al*., 2003). Thus, understanding regulation of this system will provide insight into potential targets for MTB treatment.

In *Mycobacterium smegmatis*, a model organism for MTB, the genes *esxB* and *esxA*, which encode the EsxB and EsxA proteins, are transcribed together, along with the upstream genes encoding PE35 and PPE68, proteins of unknown function (Shell *et al.*, submitted manuscript). An mRNA cleavage site exists between the regions encoding PPE68 and EsxB, allowing *esxB* and *esxA* to form a smaller, more stable transcript (Shell *et al*., submitted manuscript). After this mRNA processing event, *esxB* has a 70 nucleotide 5' untranslated region, termed the *esxB* 5' UTR, which is predicted to fold to form two hairpin loops, leaving only one free nucleotide at the 5' terminus (Shell *et al.*, submitted manuscript). Based on the previously discussed reports linking 5' stem loop structures to decreased half-lives, it is possible these hairpin loops interfere with RNase E activity, given that they are at the 5' terminus. Thus, we hypothesized that the increased stability of the *esxB-esxA* transcript post-mRNA processing is due to the formation of two hairpin loop structures at the 5' terminus of the 5' UTR.

To assess the effect of the *esxB* 5' terminal hairpin loops on mRNA stability, we created reporter constructs in which the *esxB* 5' UTR was linked to *yfp*. We introduced deletions of both or either of the 5' hairpin loops, and changes in mRNA expression, YFP protein expression, and mRNA half-life were measured. Surprisingly, deletion of both hairpin loops yielded higher mRNA levels, as indicated by quantitative PCR (qPCR), and higher YFP protein levels, as indicated by flow cytometry and microscopy. In addition, deletion of either or both loops had no significant effect on mRNA half-life, suggesting the hairpin loops at the terminus of the *esxB* 5' UTR have no stabilizing effect on the *esxB-esxA* transcript.

To assess the role of 5' UTR secondary structure in determining mRNA degradation rate on a transcriptome-wide scale, we computed base-pairing probabilities within 5' local structures using ViennaRNA and compared these to degradation pattern and half-life in MTB and *M. smegmatis*. Regions of various size starting at the 5' terminus and in the ribosome binding site of leadered transcripts or start codon of leaderless transcripts were examined for local structure predictions. We found that more structured 5' ends do not predict slower mRNA degradation in MTB or *M. smegmatis*. However, less base pairing at the ribosome binding site and start codon were found to predict slower mRNA degradation and greater ribosome occupancy in *M. smegmatis*, suggesting mRNA stability may be substantially influenced by translation efficiency in this organism.

Methods

Plasmid Construction

All plasmids were designed in Benchling based on previous constructions from the Shell Lab. The backbone comes from the plasmid pSS303 (Nguyen *et al*., 2020). It contains a Giles integration site and Hygromycin resistance marker derived from pGH1000A (Morris *et al*., 2008). It also has the reporter gene *yfp* with a C-terminal His-tag, flanked by an upstream terminator, tsynA, a synthetic downstream bi-directional terminator, ttsbi (Czyz *et al*., 2014 and

Huff *et al*., 2010, respectively), and the Pmyc1tetO promoter (Ehrt *et al*., 2005). The reporter *yfp* was replaced by an insert of the full *esxB* 5' UTR tagged downstream with a reporter *yfp* of identical sequence from a different plasmid, pSS139 (de Rivera 2016). Mutations in the *esxB* 5' UTR were made by deleting the loop regions once the full *esxB* 5' UTR insert was constructed. HiFi assembly (NEB) was used for all plasmid construction.

Primers were designed in Benchling to be about 18 nucleotides in length and annealing temperatures within 5°C of each other according to New England Biolab's T_m Calculator (Table 1). Primers for the insert sequence and deletions were designed to have an additional 18 nucleotide overlap with the backbone sequence for HiFi Assembly (New England Biolabs). The backbone and insert were amplified from the previously mentioned plasmids using PCR. Reactions (25 µL) contained 1.25 µL each of 10 µM forward and reverse primers, 0.5 µL of 10 mM each dNTPs, 1 µL of 20 ng/µL template DNA, 10.75 µL of water, 5 µL of 5X QC Enhancer, 5 µL of 5X Q5 Reaction Buffer, and 0.25 µL of Q5 Polymerase (New England Biolabs). For cycling, samples were denatured at 98°C for 30 seconds followed by denaturing at 98°C for 20 seconds, annealing according to New England Biolab's T_m Calculator, and elongation at 72°C at 30 seconds per kilobase of template for 32 cycles, followed by final elongation at 72°C for 1 minute per kilobase of template.

Table 1. Primers used in this study

Name	Sequence	Purpose
SSS1852	ATCACTAGGGCGTTGCCTCAATCGGCCTGCTGCCTG	Forward primer for pSS303 backbone
SSS1443	TCCCAGAGCCTATCTATCAC	Reverse primer for pSS303 backbone
SSS1853	GATAGATAGGCTCTGGGACTTTCCGGTGCACTCGCCGG	Forward primer for *esxB* full 5' UTR insert
SSS1377	GGCAACGCCCTAGTGATGGTGATGGTGATGAC	Reverse primer for *esxB* full 5' UTR insert
SSS1976	ATAGATAGGCTCTGGGAACTCGCCATGGAATTGGT	Forward primer *esxB* 5' UTR loop 1 deletion
SSS1984	TGCACTCGCCGGAAGAGACACAGGGAAATAAGGGGA	Forward primer *esxB* 5' UTR loop 2 deletion
SSS791	TCTTCCGGCGAGTGCACC	Reverse primer *esxB* 5' UTR loop 2 deletion
SSS1978	ATAGATAGGCTCTGGGAGACACAGGGAAATAAGGGG	Forward primer *esxB* 5' UTR loop 1&2 deletion
SSS1172	ACATATCTGTCGAAGCGCCC	Forward primer checking PCR at Giles site, left
SSS1174	TGACGATCAACTCCGCGGGGCCGGGCCA	Reverse primer checking PCR at Giles site, left
SSS1173	ACATATCTGTCGAAGCGCCC	Forward primer checking PCR at Giles site, right
SSS1175	CGGTGGATCCGCGCAACCTG	Reverse primer checking PCR at Giles site, right
SSS833	GATAGCACTGAGAGCCTGTT	Forward primer for *yfp* qPCR
SSS834	CTGAACTTGTGGCCGTTTAC	Reverse primer for *yfp* qPCR
JR273	GACTACACCAAGGGCTACAAG	Forward primer for *sigA* qPCR
JR274	TTGATCACCTCGACCATG	Reverse primer for *sig A* qPCR

The PCR products were treated directly with 0.5 µL of DpnI (New England Biolabs) at 37ºC for 15 minutes and 80ºC for 20 minutes to cleave the remaining methylated template plasmid. The products were then run on a 1% TAE agarose gel to confirm product size (Genesee). Once confirmed to be the expected size, the bands were cut out and the DNA was extracted using the Zymoclean Gel DNA Recovery Kit (Zymo). Plasmid concentrations were measured using a Nanodrop (Thermo Fisher).

Gibson Assembly was done using New England Biolab's HiFi Assembly procedure (New England Biolabs). Approximately 50 ng of vector plasmid was used with enough insert to achieve a 2:1 molar ratio based on the insert concentration, according to the New England Biolabs Ligation Calculator. Vector and insert were diluted so the previously mentioned quantities reached a final volume of 2.5 µL, and 2.5 µL of 2X HiFi DNA Assembly Master Mix (New England Biolabs) was added. The reaction was incubated at 50°C for 1 hour.

After *E. coli* transformation and sequencing (described in the section below) confirmed proper construction of the *esxB* full 5' UTR plasmid, deletions of the 5' UTR loops were made. The same plasmid construction procedure described above was used, except rather than using a backbone and an insert, the newly constructed plasmid was used as the template with the appropriately designed primers to create the desired deletions during the PCR. Additionally, for HiFi Assembly, instead of adding and calculating insert and backbone, 2.5 µL of DNA, which contained 145-215 ng, was added to 2.5 µL of Master Mix.

In addition to the constructed plasmid, all steps below from transformation into *M. smegmatis* onward also included the plasmid pSS314 with a mutated promoter followed by a reporter *yfp* gene (Nguyen 2019) to serve as the negative control in measuring *esxB* expression.

Bacterial Cultures

The HiFi assemblies were transformed into *E. coli* by adding 3 microliters of the HiFi Assembly reaction to 50 µL of NEB 5-alpha competent cells (New England Biolabs) on ice, followed by heat shock at 42ºC for 30 seconds. Then, 500 µL of SOC media were added for a 1 hr incubation in a 37ºC shaker at 200 rpm. The transformed *E. coli* were grown on LB Agar plates with 200 µg/mL of hygromycin, then cultured in LB liquid broth at 200 rpm, with 200 µg/mL of hygromycin. The plasmids were extracted from *E. coli* using the ZR Plasmid Miniprep kit (Zymo) according to the manufacturer's instructions. They were then sent to Eton Bioscience Inc. for sequencing to confirm that the desired *esxB*-5' UTR-*yfp* variations were present.

Plasmids confirmed to contain the correct mutations were transformed into *M. smegmatis* strain mc^2155. Approximately 200 ng of plasmid were added to 10 µL of competent *M. smegmatis* cells and subject to electroporation, followed by a 2.5 hr incubation at 37ºC in 200 µL 7H9. The cells were then plated on 7H9 media with 200 µg/mL hygromycin.

Colonies were checked for correct plasmid integration into the Giles site using previously documented primers in the Shell lab. Colony checking PCR (25 µL) was done using 0.2 µL of each 10 µM primer, 0.2 µL of 10 µM each dNTPs, 0.5 µL DMSO, 7.85 µL of water, 1 µL of 10X Taq Standard buffer and 0.05 µL of Taq polymerase (New England Biolabs). For cycling,

samples were denatured at 95°C for 5 minutes followed by 32 cycles of denaturing at 95°C, annealing at 56°C, and elongation at 65°C. The PCR products were run on a 1% TAE agarose gel to confirm band size (Genesee). Confirmed colonies were grown in liquid 7H9 in a 37°C 200 rpm shaker until turbid, then stored in 25% glycerol stocks.

Fluorescence Microscopy

M. smegmatis stocks were thawed and 50 µL of each were added to 5 mL of 7H9 media. The cultures were grown overnight, then normalized to grow to an OD 0.8 the next day. A large glass slide was sprayed with water repellent and let to dry. A 50 µL drop of 1% agar was added to the repellent surface, and a microscope slide was immediately pressed onto the drop and held to dry. Without concentrating the cells, the cultures were vortexed for resuspension. A 5 µL drop of sample was added to the center of the agar pad and a cover slip was placed on top. The samples were imaged with a Zeiss Axio Imager Z1 with the Zeiss Apotome (Zeiss) under 40X oil immersion. Cells were imaged using a DIC 40 channel and YFP channel for brightfield and fluorescent images. Identical display parameters were used for each image.

Flow Cytometry

M. smegmatis stocks were thawed and 50 µL of each were added to 5 mL of 7H9 media. The cultures were grown overnight, then normalized to grow to an OD 0.6 the next day in freshly filtered media to eliminate precipitates. For the rest of the protocol, samples were kept on ice to prevent overgrowth. Samples were diluted to an OD of 0.015 with 7H9 and filtered through a 5 µm filter needle to remove any clumps. Following controls with water and 7H9 to set appropriate thresholds for background reduction, the samples were run on a CytoFLEX flow cytometer (Beckman Coulter Life Sciences). For each sample, 60,000 events were collected, and an appropriate gate was set for comparison between samples.

RNA Extraction

M. smegmatis stocks were thawed and 50 µL of each were added to 5 mL of 7H9 media. The cultures were grown overnight, then normalized to grow to an OD 0.6 the next day. Once the cultures reached an OD ~0.6, they were frozen in liquid nitrogen. The following day, RNA was extracted from the frozen *M. smegmatis* cultures using the Direct-zol RNA extraction and purification kit (Zymo). The liquid cultures were thawed and then centrifuged at 4°C and 3,900 rpm for 5 minutes. Following discard of the supernatant, samples were taken to a fume hood to resuspend in 1 mL of TRIzol (Thermo Fisher) and were immediately transferred to tubes of 100 µm zirconium lysing matrix beads (OPS Diagnostics). The samples were lysed in a FastPrep 5G at 7 m/s for 30 seconds for a total of 3 cycles, with 2 minutes on ice in between each cycle (MP Biomedicals). The samples were taken back to the fume hood and 300 µL of chloroform was added to all samples.

All equipment and the bench were wiped with RNaseZAP (Thermo Fisher) to prevent RNA degradation. The samples were vortexed for 15 seconds, then centrifuged at 15,000 rpm and 4°C for 15 minutes to allow for phase separation. Without disturbing the bottom phase, up to 500µL of the aqueous phase was removed and evenly mixed with 500 µL of 100% ethanol. The samples

were transferred to Direct-Zol columns and centrifuged at 15,000 g at room temperature for 30 seconds. Aspiration with a pipet tip was used to carefully remove and discard the flow through. The samples were then washed with 400 µL of RNA Wash Buffer (Zymo). This was followed by a 1 hr DNase treatment at room temperature using 80 µL of DNase Master mix, made from 75 µL of DNase Digestion Buffer and 5 µL of DNase I (Zymo). The samples were then washed twice with 400 µL of Direct-Zol Pre-wash Buffer and a 30 second centrifugation, followed by a final wash of 700 µL of RNA Wash Buffer with a 2 minute centrifugation (Zymo). The Direct-Zol columns were placed in a clean 1.5 mL tube and the RNA was eluted in 50 µL of RNase-free water with a 1 minute centrifugation. The extracted RNA was resuspended and measured for concentration and purity using a Nanodrop (Thermo Fisher).

cDNA Synthesis and Clean Up

To prevent RNA degradation, the equipment and bench were wiped with RNaseZAP (Thermo Fisher). RNA samples were thawed, and 600 ng of each RNA sample was added to water for a final volume of 5.25 µL. This was repeated to have two tubes of each diluted RNA, one for reverse-transcription (RT) one for no-RT controls. A random primer master mix (1X) was made from 0.5 µL Tris, 0.33 µL of water, and 0.17 µL of 3 µg/µL random primer mix (New England Biolabs), scaled according to the number of samples, and 1µL of this mix was added to each sample. The samples were incubated at 70°C for 10 minutes, then placed in on ice for 5 minutes.

RT master mix contained 0.5 µL of 10 mM dNTPs, 0.5 µL of 100 mM DTT, 0.25 µL murine RNase inhibitor, 2 µL 5X ProtoScript Buffer, and 0.5 µL of ProtoScript II (New England Biolabs) per reaction. No-RT master mix was made the same way, except water was added instead of ProtoScript II. 3.75 µL of RT master mix was added to one of the two RNA tubes for each sample, while 3.75 µL of no-RT master mix was added to the other tube for each sample. The samples were incubated at 25°C for 10 minutes and at 42°C for 2 hours, then 4°C overnight.

RNA was degraded from each sample with 5 µL of 0.5 mM EDTA and 5 µL of 1 N NaOH. The samples were incubated at 65°C for 15 minutes. RNA degradation was stopped by adding 12.5 µL of 1 M Tris-HCl pH 7.5 to each sample. The Monarch PCR & DNA Purification Kit (New England Biolabs) was used for cDNA clean-up. All centrifugations were done at 17,900 g for 30 seconds and flow-through was discarded by aspiration with a pipet tip. Samples were transferred to a 1.5 µL tube containing 325 µL of binding buffer, then transferred to a Monarch column and centrifuged. For a total of 3 times, 200 µL of wash buffer was added to the columns and they were centrifuged. The columns were transferred to fresh 1.5 mL tubes and 35 µL RNase-free water was used to elute the cDNA. Concentration and purity were measured on a Nanodrop (Thermo Fisher). All cDNA samples were stored at -20°C.

Quantitative PCR (qPCR)

RT cDNA samples were diluted to 1 ng/µL using RNase-free water and further diluted to bring the final concentration to 200 pg. For the no-RT samples, water was added in amounts equal to the corresponding RT sample. A 96 well plate was designed to measure the expression of *yfp* for RT samples only, and sigA, which will serve as a control for expression, for all RT and no-RT samples. In each well, 2 µL of the appropriate sample was added. Primer mixes for sigA and *yfp*

were made to contain 2.5 μM of each primer as indicated in Table 1. qPCR master mixes (1X) were made using 1 μL of 2.5 μM primer mix, 5 μL of iTaq SYBR Green Supermix (BioRad), and 2 μL of RNase-Free Water. 8 μL of the appropriate qPCR master mix was added to each well containing 2 μL cDNA. The plate was sealed and placed in an Applied Biosystems 7500 Real-Time PCR System. The plate was incubated at 50°C for 2 minutes, 95°C for 10 minutes, and 40 cycles of 95°C for 15 seconds and 61°C for 1 minute. Once the run was complete, the threshold was changed to 0.2, and the data were exported.

Half-life Measurement

M. smegmatis stocks were thawed and 50 μL of each were added to 5 mL of 7H9 media. The cultures were grown overnight and normalized to grow to an OD 0.7 in 10 mL of 7H9 media the next day to confirm consistent growth between cultures. The cultures were normalized to grow to an OD of 0.8 in 50 mL of 7H9 media in 250 mL culture flasks the next day. In a 37ºC warm room, each culture was split into five 8-mL cultures and placed on a culture wheel for 30 minutes. Rifampin was added to the cultures at a concentration of 150 ug/mL and they were frozen in liquid nitrogen after 0 seconds, 30 seconds, 1 minute, 2 minutes, or 4 minutes. RNA extraction, cDNA synthesis and cleanup, and qPCR were done following the exact procedures described above for all timepoints and replicates of each strain.

Bioinformatics Data Collection

Fasta files composed of gene IDs and mRNA sequences of annotated genes in *M. smegmatis* (Martini *et al.,* 2019) and MTB (Shell *et al.,* 2015) were obtained from within the lab. Lab members created these based on annotated transcription start sites. Genes that are either first in an operon or monocistronic were selected for analysis (Appendix A). This included both leadered genes and leaderless genes, which were considered separately. Regions of coding sequences and 5' UTR sequences of various length were extracted using Python scripting and stored in separate files for batch analysis. Table 2 summarizes the regions of mRNA sequences chosen for analysis. These sequence regions were extracted for both MTB and *M. smegmatis*.

For each of these sequence regions in both species, minimum free energy and unpaired probabilities within the predicted secondary structure were calculated as detailed below and stored in Excel based on gene ID, along with the gene's half-life. MTB half-lives reported by Rustad *et al.* were obtained from the supplementary data section of the report (Rustad *et al.,* 2013). *M. smegmatis* half-life calls were made by Professor Scarlet Shell for the majority of the genome using data collected within the Shell lab. Transcripts were also clustered into slow (S), medium (M), and fast (F) decay groups based on degradation profiles by Huaming Sun of the Shell Lab. This was used to assess the pattern of decay in relationship to secondary structure, and was useful in providing an additional layer of information on top of the half-life calls.

Ribosome occupancy was also compared to half-life and secondary structures in *M. smegmatis* transcripts. The original data was reported as supplementary material by Chen et. al and occupancy at regions of interest were determined within the Shell lab. In particular, ribosome

occupancy at the region that covers the last 20 nucleotides of the 5' UTR plus the first 18 nucleotides of the coding sequence was analyzed.

Secondary Structure Prediction

Secondary structures were predicted using the ViennaRNA package (Lorenz *et al*. 2011). The RNAfold command was used to compute the most likely structure of the mRNAs based on minimizing the free energy. A Python script was written and used to export the gene ID and minimum free energy (MFE) of each predicted structure from the resulting text file into an Excel spreadsheet.

The RNAplfold command was used to compute unpaired probabilities for a stretch of nucleotides in the transcript with the input parameters u set to 5 and W set to the length of the input sequence. For analysis of full 5' UTRs, a Python script was written for dynamic analysis so that the files could be run in batch with the W parameter set to the length of each individual sequence. This allows each nucleotide to base pair with any other nucleotide in the sequence. The third number in the third column of the resulting matrix file for each sequence was extracted and added to the Excel file. This number represents the probability that all of the first 3 nucleotides in the sequence are unpaired. The same was done for the fifth number in the fifth column to find the probability that all of the first 5 nucleotides are unpaired.

To represent pairing probabilities in ribosome binding regions, an average was taken of the probability that each nucleotide and the previous nucleotide are both unpaired. This was done for the last 30 nucleotides of the 5' UTR plus either the first 20 nucleotides of the coding sequence or the start codon.

Statistical Analysis

Data analysis was done in Python using the Pandas, NumPy, and SciPy libraries to import data, calculate correlations, and produce plots. Spearman's correlation was calculated between the MFE of each portion of the sequence and that transcript's half-life. Since the distribution of the unpaired probabilities was bimodal, the data was organized into 5 bins of size 0.2 based on probability. Bee-swarm plots with overlayed boxplots were made for each of the bins. A Kruskal Wallis analysis was done to determine the relationship between the median value of each bin. If the p-value of the analysis was less than or equal to 0.05, Dunn's post-test was performed to determine which bins had a significant difference.

Results

Construction of reporter plasmids to test the impact of secondary structure in the *esxB* 5' UTR

To test the effects of the 5' hairpin loops in the *esxB* 5' UTR, four plasmids were constructed in which variants of the *esxB* 5' UTR were linked to sequence encoding the fluorescent protein YFP (Figure 1). One plasmid was constructed to contain the full wildtype *esxB* 5' UTR. Two plasmids were constructed with one loop deletion: one contains deletion of the first 23 nucleotides to remove the first loop and one contains deletion of nucleotides 25-44 relative to the 5' end to remove the second loop. It is of note that the latter two constructs were designed to have one unpaired nucleotide at the 5' end since it is required for RppH and RNase E activity in *E. coli* (Deana *et al*., 2008), and to match the single unpaired nucleotide at the 5' end of the full *esxB* 5' UTR for consistency. The fourth construct has the first 44 nucleotides deleted to remove both loops, resulting in a long stretch of unpaired nucleotides at the 5' end. In addition, an existing plasmid in the Shell lab with the same backbone and YFP reporter as our constructs, but a mutated Pmyc1tetO promoter with a proven lack of YFP expression, was transformed to serve as a negative control for the expression-based experiments. The secondary structures of the UTR variants were predicted using ViennaRNA Package and show the expected loop deletions (Figure 2).

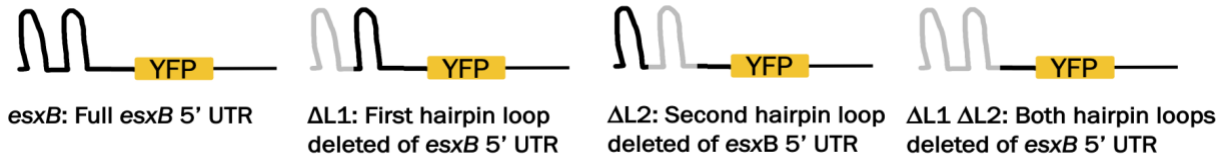


Figure 1. Plasmid inserts were made with four variations of the *esxB* 5' UTR with the reporter *yfp*. All plasmids had the same promoter upstream of the UTR (not shown). The expression cassettes shown were placed in a plasmid backbone that integrates into the genome as a single copy.

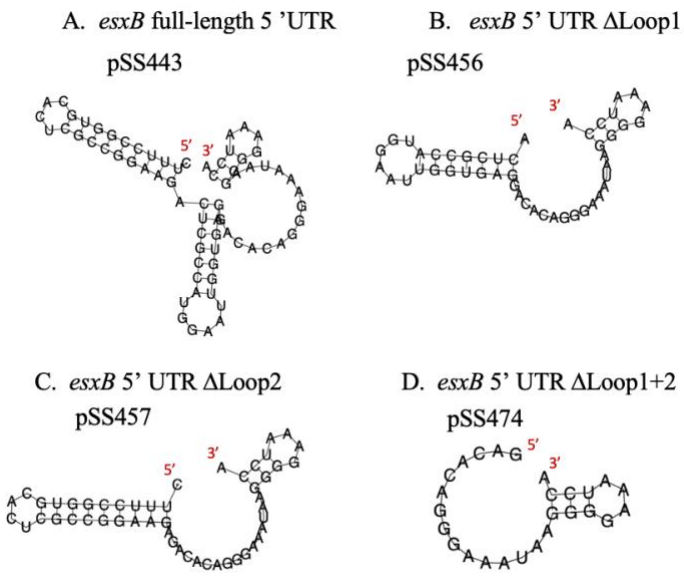


Figure 2. Predicted secondary structures of *esxB* RNA 5' UTR loop deletions.
(A) The *esxB* 5' UTR is predicted to have two hairpin loops at the 5' end. (B) Deletion of the first 23 nucleotides removes the first hairpin loop. (C) Deletion of nucleotides 25-44 relative to the 5' end removes the second hairpin loop. (D) Deletion of the first 44 nucleotides removes both hairpin loops of the *esxB* 5' UTR. Structures were predicted using ViennaRNA Package.

<u>Deletion of both *esxB* 5' UTR hairpin loops increases mRNA levels, while deletion of just one loop has no effect on mRNA levels</u>

To assess the effect of the *esxB* 5' hairpin loops on mRNA abundance, qPCR was done to determine *yfp* expression levels relative to levels of the housekeeping gene sigA (Figure 3). We predicted that deletion of both terminal loops of the *esxB* 5' UTR would cause decreased mRNA stability and therefore decreased steady-state abundance, while deletion of just one loop would cause little decrease or no effect. Indeed, no statistically significant difference was observed in *yfp* expression with deletion of either loop relative to the full *esxB* 5' UTR. However, surprisingly, significantly increased mRNA abundance was observed upon deletion of both hairpin loops of the *esxB* 5' UTR (Figure 3).

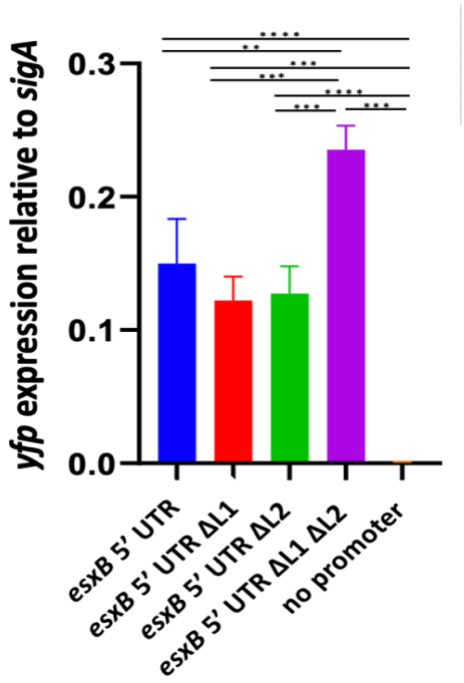


Figure 3. qPCR analysis shows the effects of *esxB* 5' UTR loop deletions on *yfp* mRNA expression. Triplicate cultures for each strain were grown to log phase. mRNA expression of *yfp* is normalized to the housekeeping gene *sigA*. Bars show mean expression, and error bars show standard deviation. Significance was calculated by a one-way ANOVA and post-hoc Tukey's test. ** P < 0.01; ***, P < 0.001; ****, P < 0.0001. Comparisons that were not statistically significant are not shown.

YFP protein levels are increased by deletion of both *esxB* 5' UTR hairpin loops and decreased by deletion of either loop

To test the effect of the *esxB* 5' hairpin loops on protein abundance, microscopy and flow cytometry were done to determine YFP expression levels. Microscopy results show a noticeable increase in YFP fluorescence with the deletion of both *esxB* 5' UTR loops compared to the wildtype UTR, while deletion of either loop had no noticeable effect on YFP expression (Figure 4).

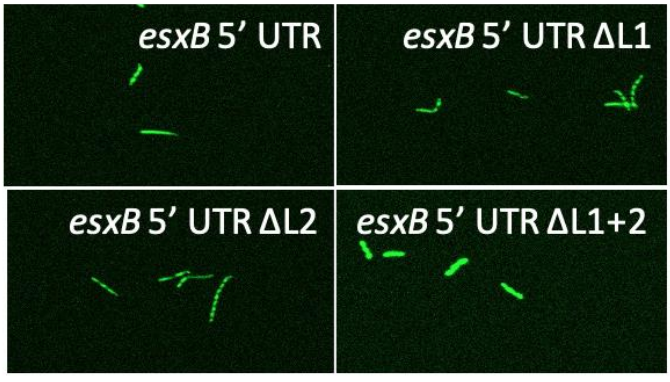


Figure 4. Fluorescence microscopy shows the effect of *esxB* 5' UTR loop deletions on YFP protein levels. Triplicate cultures in log phase growth were imaged under 40X oil immersion. The same exposure and display settings were used for all strains.

Flow cytometry was then used to quantify YFP protein level. To compare similarly sized cells of each strain, a single gate was created to capture the densest portion of each forward-scatter versus side-scatter plot and applied to each strain (Figure 5). The gated region contained 3,000-3,500 cells for each strain. Analysis of the cells within the gate showed that deletion of either the first or second loop of the *esxB* 5' UTR caused a decrease in median YFP protein expression (Figure 6). However, higher median YFP expression was observed when both hairpin loops were deleted compared to the full-length UTR (Figure 6). The impact of the loops therefore seems to differ depending on whether they are present individually or in combination.

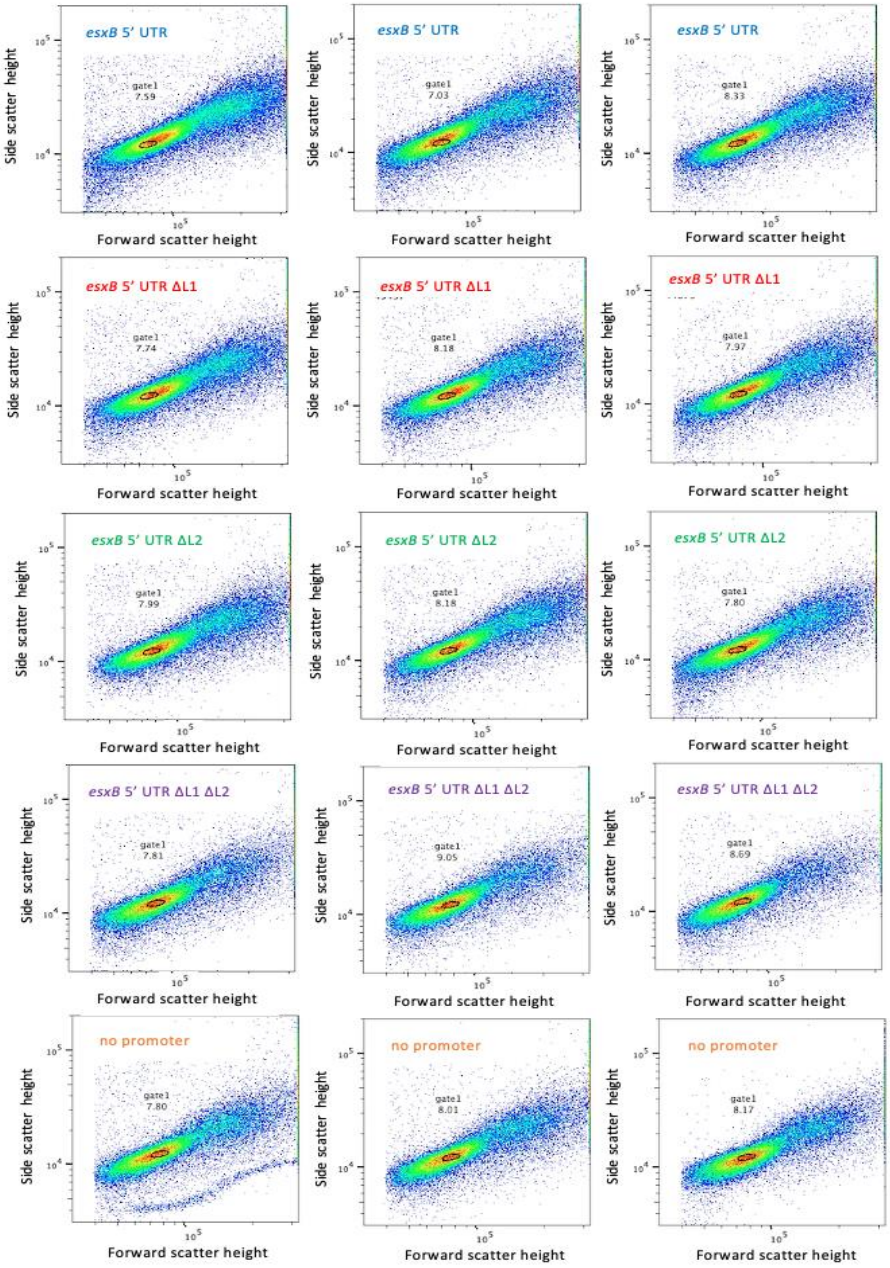


Figure 5. Gate chosen for flow cytometry analysis.
A gate was created (black oval) and applied to each forward scatter versus side scatter plot to ensure that similarly sized cells were used for comparisons between strains. Data from each of three replicates per strain are shown.

A.

Strain	Median Fluorescence	Coefficient of Variation
esxB 5' UTR	1288	0.20
	1293	0.24
	1263	0.22
esxB 5' UTR ΔL1	673	0.28
	750	0.24
	775	0.24
esxB 5' UTR ΔL2	742	0.25
	814	0.23
	840	0.25
esxB 5' UTR ΔL1 ΔL2	2982	0.24
	2987	0.22
	3047	0.22
no promoter	70.5	0.85
	69.1	0.84
	70.3	0.87

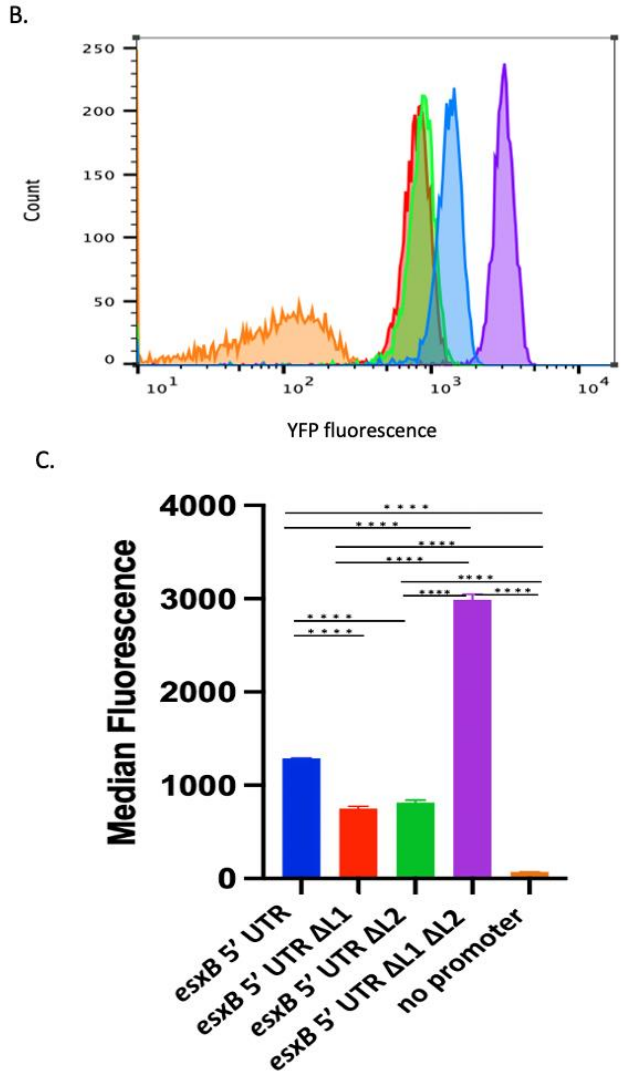


Figure 6. Flow cytometry shows the effects of *esxB* 5' UTR loop deletions on YFP protein levels.
Flow cytometry was done with triplicate cultures of each strain in log phase growth. (A) Median YFP fluorescence and coefficient of variation were measured for each gated region. (B, C) The second replicate of each strain (highlighted) were chosen for the analysis plots since there was consistency between replicates. (C) Bars show median fluorescence, and error bars show 95% CI. Significance was found by a Kruskal-Wallis non-parametric test, followed by a Dunn's post-hoc test. ****, P < 0.0001.

mRNA half-life of *yfp* is not affected by *esxB* 5' UTR hairpin loop deletions

To test the effects of the 5' hairpin loops in the *esxB* 5' UTR on mRNA degradation, the half-life of transcripts with deletions of the loops was determined. Results of the experiment show no significant difference in mRNA half-life in transcripts with any *esxB* 5' loop deletions compared to the full-length UTR (Figure 7 A&B). It is of note that there appears to be some noise in the data (Figure 7A), resulting in large 95% confidence intervals. We therefore cannot exclude the

possibility that there are small differences in half-life that we were unable to detect. Nonetheless, our data collectively allow us to draw some conclusions regarding the effects of the *esxB* 5' UTR hairpin loops.

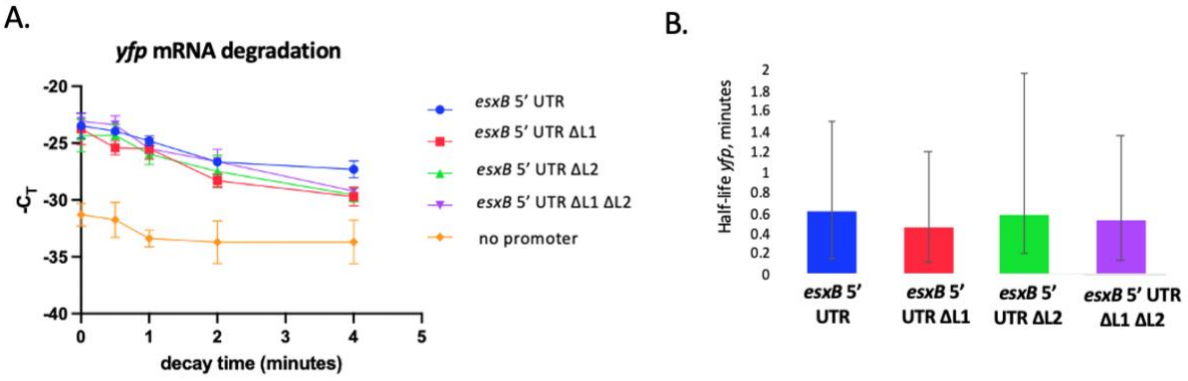


Figure 7. mRNA half-life analysis shows the effects of *esxB* 5' UTR loop deletions on mRNA degradation rate. Triplicate cultures of each strain were grown to log phase and rifampicin was added to halt transcription. Cultures were frozen 0 seconds, 30 seconds, 1 minute, 2 minutes, and 4 minutes after rifampicin addition. (A) qPCR was done to analyze mRNA levels at each time point. The y axis uses the -CT because this reflects relative abundance on a log2 scale. (B) mRNA degradation rates were compared using linear regression and half-lives were determined by the negative reciprocal of the best-fit slope for the time-points between 0 and 2 minutes, shown by the bars. Error bars show 95% confidence intervals.

Overall, these experiments show that the *esxB* 5' UTR hairpin loops do not substantially stabilize the *yfp* mRNA. If the loops were stabilizing the transcript, as initially hypothesized, we would expect deletion of the loops to decrease *yfp* half-life and mRNA and protein abundance. Surprisingly, deletion of both 5' loops increased both *yfp* mRNA and protein expression levels relative to the full UTR. This differs from the expectation that eliminating base pairing at the 5' end will enable RNase E to more efficiently degrade the transcript (Deana *et al.*, 2008) (see Discussion).

Unpaired 5' ends do not predict mRNA stability transcriptome-wide

To assess the relationship between base-pairing near the 5' ends of transcripts and mRNA stability, the probability of unpaired bases in various regions of the transcripts were compared to half-life measurements. For this analysis, sequences of transcripts for genes in *M. smegmatis* and MTB that are either first in an operon or monocistronic were input to ViennaRNA secondary structure prediction software using the "RNAplfold" command (Lorenz *et al.* 2011). This group includes both leadered and leaderless transcripts, which were analyzed separately since the 5' end is also the ribosome binding site (RBS) in leaderless transcripts. A full list of the genes included in the analysis can be found in Appendix A.

In addition to assessing structures within full-length 5' UTRs, small portions of the sequences were input to the software to account for the fact that the first few nucleotides may have more opportunities to base-pair with nearby nucleotides than with those further away, such as at the distal ends of long UTRs. Regions of the coding sequence were assessed in leaderless transcripts,

and the RBS, including the beginning of the coding sequence, was assessed in leadered transcripts. A full summary of the regions input to ViennaRNA are summarized in Table 2 below. From the ViennaRNA output, numbers which represent the probability that the first three or five nucleotides are unpaired were calculated for each transcript and compared to its half-life. Half-lives for MTB genes were obtained from published work (Rustad *et. al* 2013) and half-lives for *M. smegmatis* genes were available from other ongoing work in our lab.

Table 2. Regions of mRNAs subjected to secondary structure analysis for leadered and leaderless genes in both *M. smegmatis* and MTB. nts, nucleotides; and CDS, coding sequence.

Sequence Region Allowed to Fold	Gene Type	Identifier
Full 5' UTR	Leadered, *M. smegmatis* & MTB	Full UTR 
First 20 nts 5' UTR	Leadered, *M. smegmatis* & MTB	First 20 UTR 
First 30 nts 5' UTR	Leadered, *M. smegmatis* & MTB	First 30 UTR 
First 40 nts 5' UTR	Leadered, *M. smegmatis* & MTB	First 40 UTR 
Last 30 nts 5' UTR + first 20 nts CDS	Leadered, *M. smegmatis* & MTB	Last 30 UTR +First 20 CDS 
Last 30 nts 5' UTR + start codon	Leadered, *M. smegmatis* & MTB	Last 30 UTR + Start 
First 20 nts CDS	Leadered & **Leaderless**, *M. Smegmatis* & MTB	First 20 CDS
First 30 nts CDS	**Leaderless**, *M. smegmatis* & MTB	First 30 CDS 
First 40 nts CDS	**Leaderless**, *M. smegmatis* & MTB	First 40 CDS 

Interestingly, the distribution of pairing probabilities was bimodal (Figure 8). We therefore placed transcripts into bins according to pairing probabilities and compared the half-lives across bins. We initially predicted that transcripts with greater probabilities of having an unpaired 5' ends would degrade faster, and thus have shorter half-lives. First, the full 5' UTR and the first twenty nucleotides of the 5' UTR (leadered), or the coding sequence (CDS) (leaderless) were computationally folded and binned according to the probability that the first five nucleotides were all unpaired. Surprisingly, half-life did not differ significantly between the bins in *M. smegmatis* (Figure 9 A-C) or MTB (Figure 9 D-F). Next, the first 30 or 40 nucleotides of the 5' UTR (leadered) or CDS (leaderless) were computationally folded and binned according to the probability that the first five nucleotides were unpaired. Again, there was no significant difference in half-life between the bins in *M. smegmatis* (Figure 10 A-D) or MTB (Figure 10 E-H), suggesting the previous result is not from incorrect representation of the length of sequence which may fold in the cell.

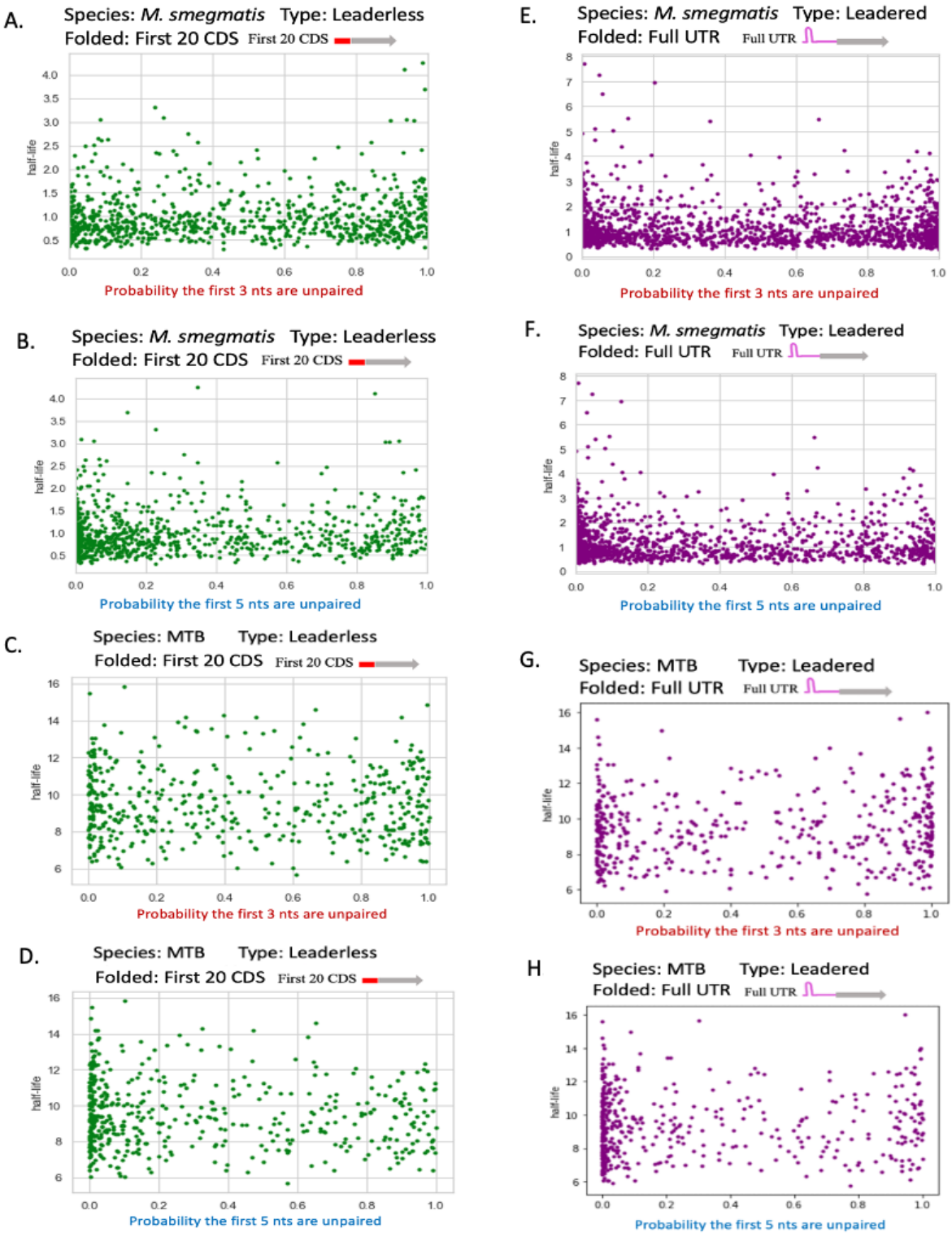


Figure 8. Relationship Between Unpaired Probabilities at 5' Ends and Half-life Shown as Scatterplots. The minimum free energy secondary structures for full 5' UTRs of leadered transcripts and the first 20 nucleotides (nts) of the coding sequence (CDS) of leaderless transcripts were calculated in Vienna RNA. Unpaired probabilities of the first 3 or 5 nucleotides calculated by the software were compared to half-life. Spearman's correlation was computed for statistical analysis, and no significance was found.

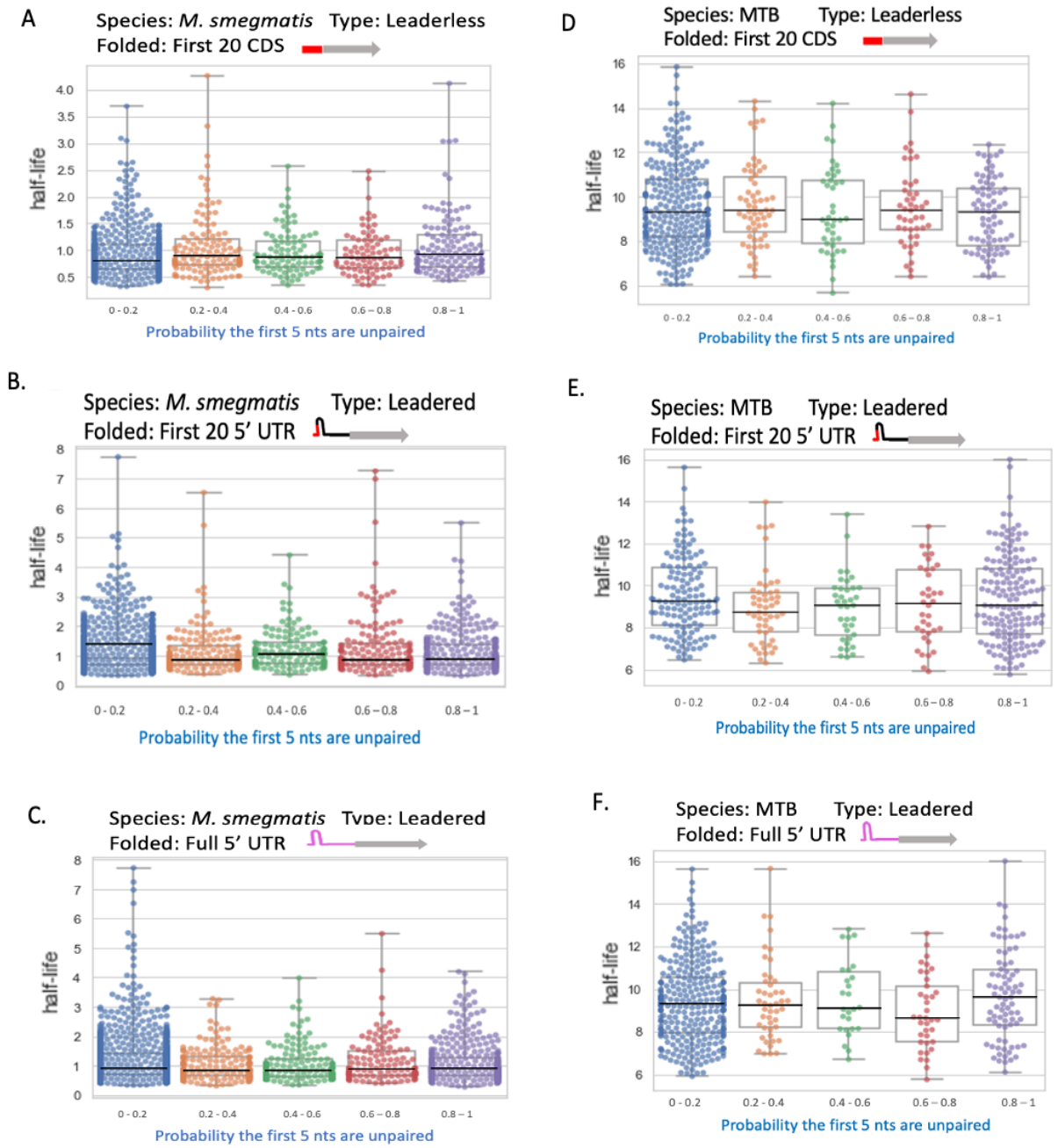


Figure 9. mRNA half-lives of transcripts binned according to the probability that the first 5 nucleotides are unpaired if the first 20 nucleotides from the 5' end or the full 5' UTR is allowed to fold.
Local MFE secondary structures of mRNAs were predicted in (A, D) for the first 20 nucleotides of the CDS (leaderless) and (B, E) the first 20 nucleotides or (C, F) full-length 5' UTR (leadered) using ViennaRNA. Analysis was done for (A-C) *M. smegmatis* and (D-F) MTB. The transcripts were binned according to probability that the first 5 nucleotides of this region were unpaired, and the median half-lives (represented by black bars) were compared across bins using a Kruskal-Wallis non-parametric test. No significant differences in half-life between the bins was found by this test.

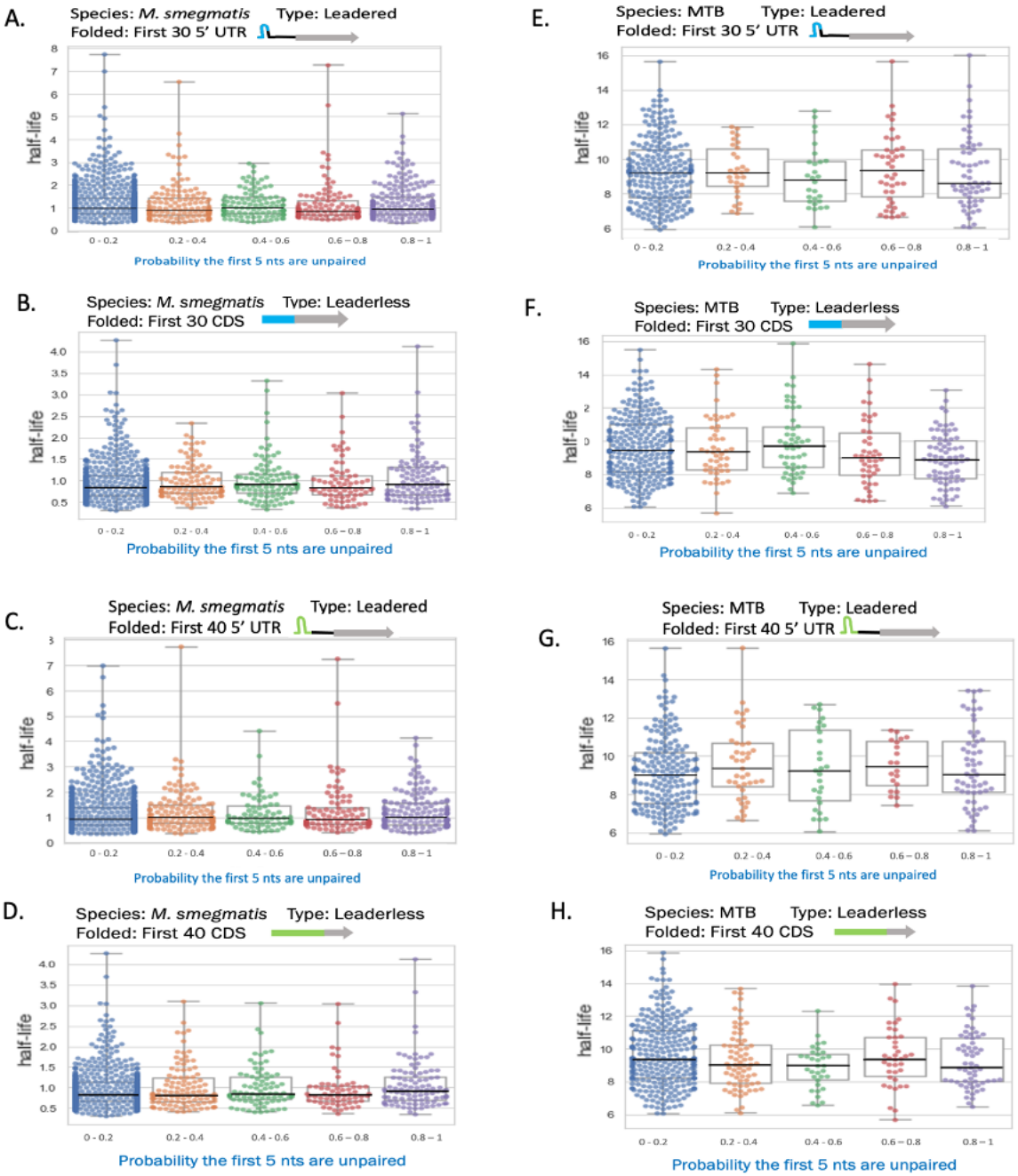


Figure 10. mRNA half-lives of transcripts binned according to the probabilities that the first 5 nucleotides are unpaired if the first 30 or 40 nucleotides at the 5' end of the sequence are allowed to fold.
Local MFE secondary structures of mRNAs were predicted in the first 30 nucleotides of the (A&E) 5' UTR (leadered) or the (B&F) CDS (leaderless) and the first 40 nucleotides of the (C&G) 5' UTR (leadered) or the (D&H) CDS (leaderless) using ViennaRNA. Analysis was done for (A-D) *M. smegmatis* and (E-G) MTB. The transcripts were binned according to probability that the first 5 nucleotides of this region were unpaired, and the median half-life values (represented by black bars) in each bin were compared using a Kruskal-Wallis non-parametric test. No significant differences in half-life between the bins were found by this test.

To determine whether a smaller number of unpaired nucleotides may predict mRNA degradation rates in mycobacteria, we assessed the probability that the first three nucleotides were unpaired. When the full 5' UTR and the first twenty nucleotides of the 5' UTR (leadered), or the CDS (leaderless) were folded and binned according to the probability that the first three nucleotides are all unpaired, there were no significant differences in half-life across bins in *M. smegmatis* (Figure 11A-C) or MTB (Figure 11 D-F). When the same analysis was done on the first 30 or 40 nucleotides of the 5' UTR (leadered) or CDS (leaderless), similarly there was no difference in half-life across bins in *M. smegmatis* (Figure 12A-D) or MTB (Figure 12 E-H).

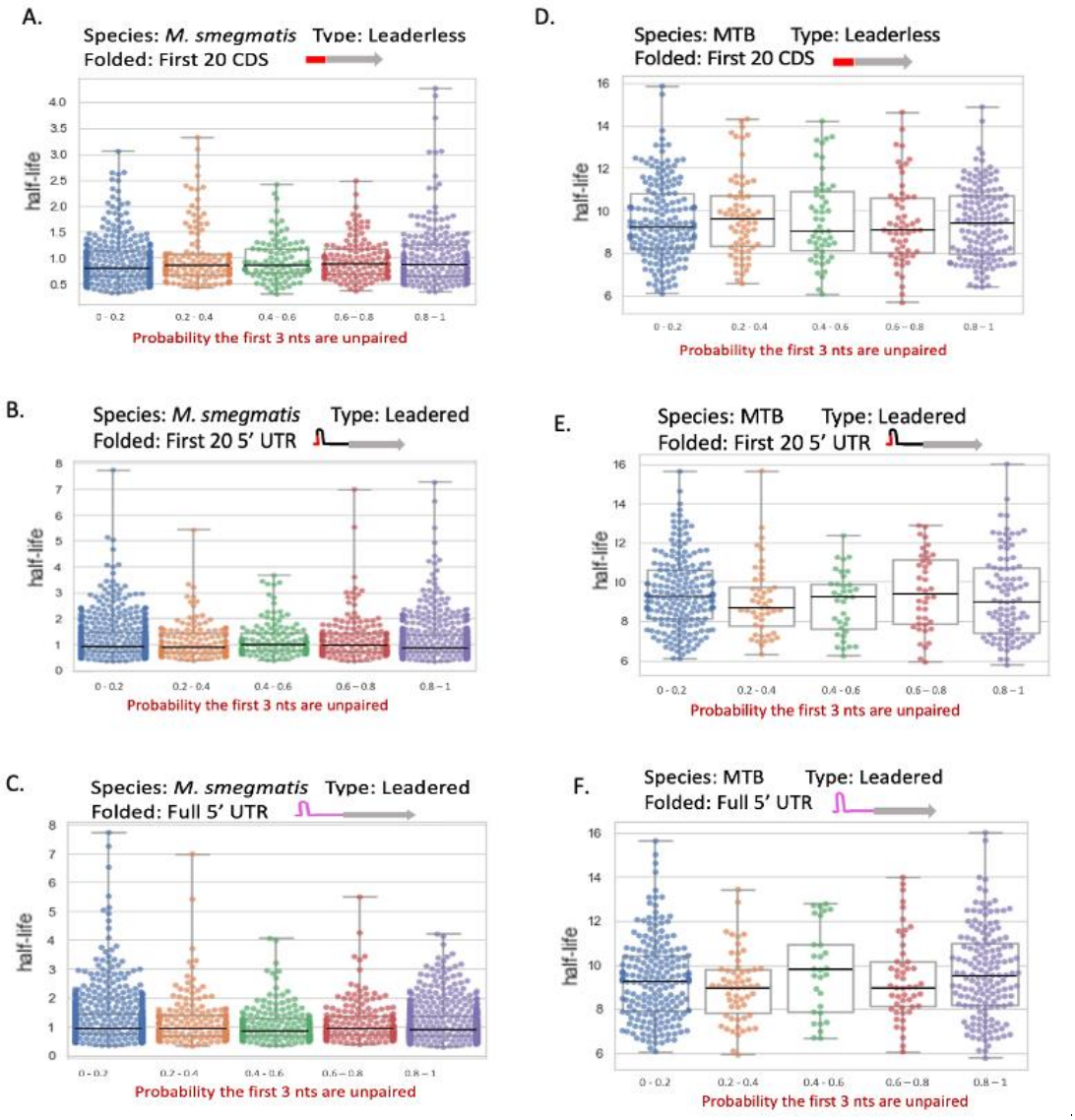


Figure 11. mRNA half-lives of transcripts binned according to the probabilities that the first 3 nucleotides are unpaired if the first 20 nucleotides from the 5' end of the sequence or the full 5' UTR is allowed to fold. Local MFE secondary structures of mRNAs were predicted in the first 20 nucleotides of the (A, D) CDS (leaderless) or (B, E) 5' UTR (leadered) or (C, F) the full-length 5' UTR using ViennaRNA. Analysis was done for (A-C) *M. smegmatis* and (D-F) MTB. The transcripts were binned according to probability that the first 3 nucleotides of this region were unpaired, and the median half-life values within the bins (represented by black bars) were compared using a Kruskal-Wallis non-parametric test. No significant in half-life between the bins was found by this test.

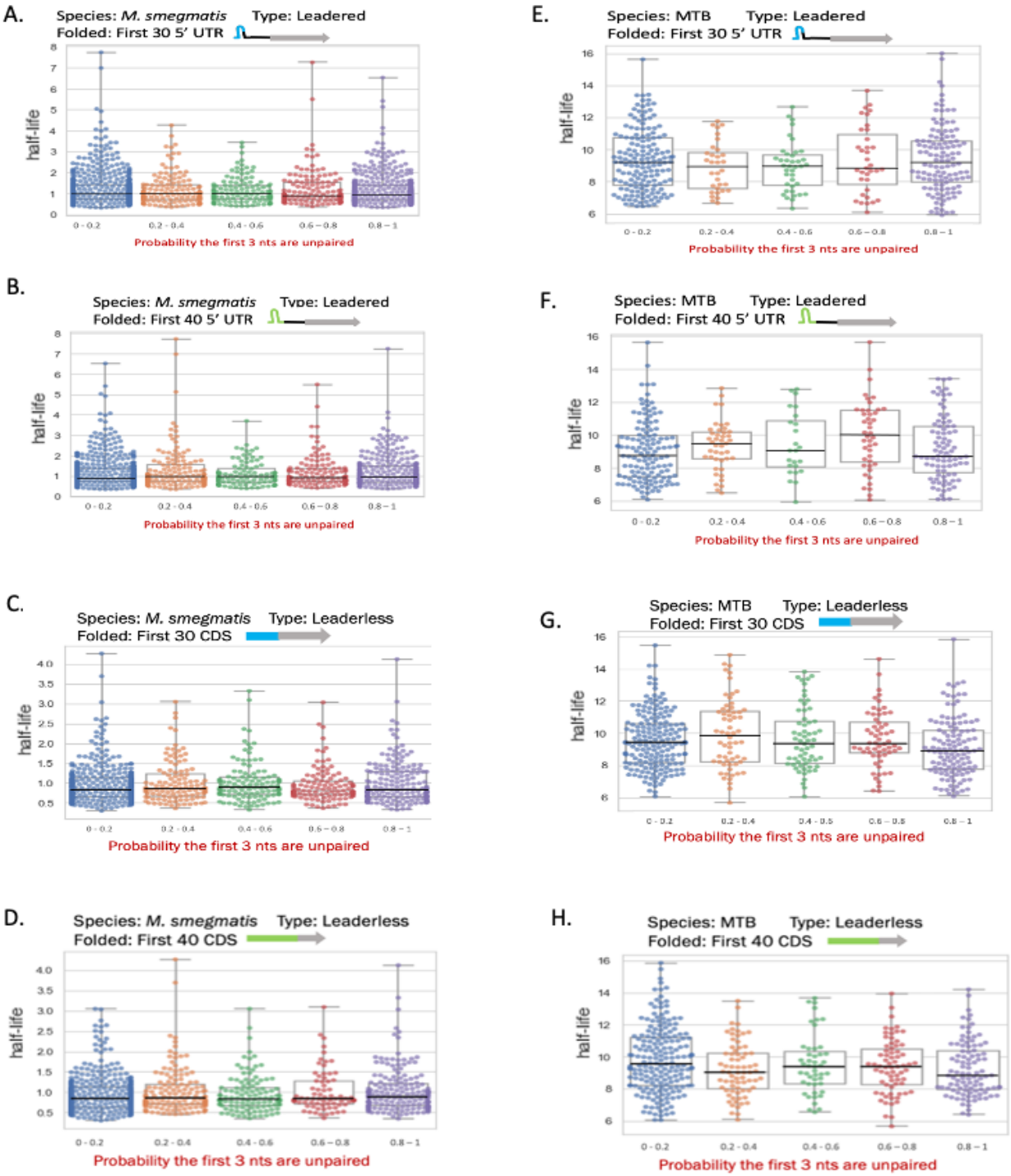


Figure 12. mRNA half-lives of transcripts binned according to the probabilities that the first 3 nucleotides are unpaired if the first 30 or 40 nucleotides from the 5' end of the sequence is allowed to fold.
Local MFE secondary structures of mRNAs were predicted in the first 30 nucleotides of the (A&E) 5' UTR (leadered) or the (B&F) CDS (leaderless) and the first 40 nucleotides of the (C&G) 5' UTR (leadered) or the (D&H) CDS (leaderless) using ViennaRNA. Analysis was done for (A-D) *M. smegmatis* and (E-G) MTB. The transcripts were binned according to probability that the first 3 nucleotides of this region were unpaired, and the median half-life values within the bins (represented by black bars) were compared using a Kruskal-Wallis non-parametric test. No significant in half-life between the bins was found by this test.

Using data from within the Shell lab which clustered *M. smegmatis* transcripts based on having slow, medium, or fast decay, we next investigated whether the probability of having 3 or 5 unpaired nucleotides from the 5' end differed between transcripts with relatively slow, medium, or fast decay rates (See Appendix B). This analysis was done using the same data set of pairing probabilities at the 5' end of transcripts in *M. smegmatis* as the half-life analysis above. Similar to the results of the half-life analysis, the probability of having unstructured 5' ends did not differ significantly between groups of slow, medium, or fast decaying transcripts.

Start codon identity influences leaderless mRNA steady-state abundance

Since the first three nucleotides of leaderless transcripts are also the start codon, we wondered if start codon identity in leaderless transcripts influenced mRNA pairing probabilities or degradation rates. Both AUG and GUG start codons are common in mycobacteria. Comparison of mRNA half-lives between transcripts with different start codon identities in *M. smegmatis* revealed no significant difference between a transcript's start codon and half-life (Figure 13A&C), leading us to conclude that the bimodal distribution of pairing probabilities is not likely due to differences in leaderless start codon. Leaderless transcripts starting with the GUG codon were on average more abundant than those starting with AUG (Figure 13D), but this difference was not evident for leadered transcripts (Figure 13B).

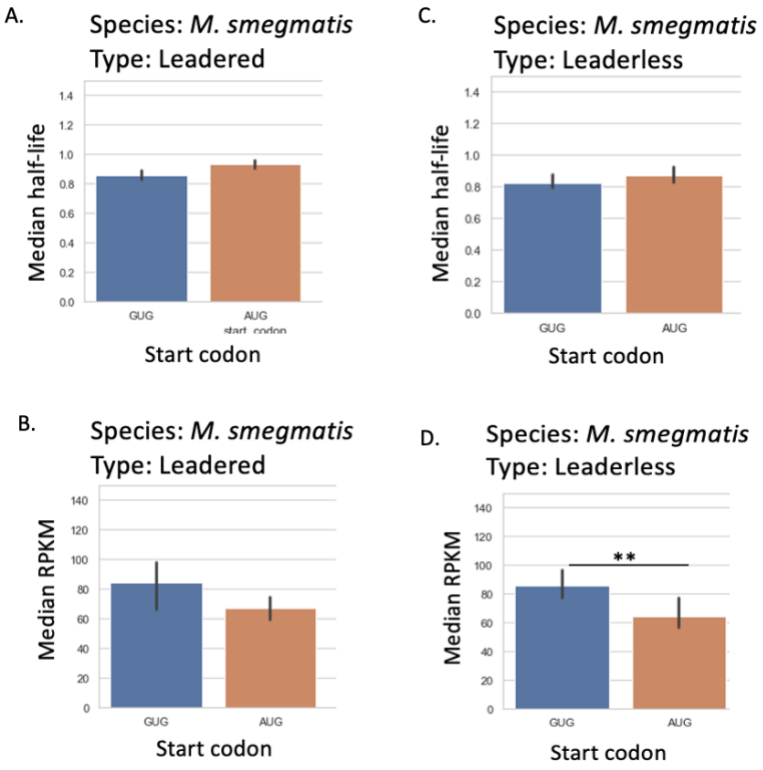


Figure 13. Relationships between start codon identities, mRNA half-life, and steady-state abundance in *M. smegmatis*.
(A&B) Leadered and (C&D) leaderless transcripts of *M. smegmatis* were placed into groups based on their start codon identity. Median values of (A&C) half-life and (B&D) steady-state abundance (RPKM) were compared by Mann-Whitney one-way analysis of variance. Error bars show the 95% confidence interval. **, $P < 0.01$.

Lower base-pairing probability in ribosome binding sites (RBSs) and start codons is associated with slower mRNA degradation and higher ribosome occupancy in some situations

We then hypothesized that secondary structures in ribosome binding sites (RBS) may be predictive of mRNA half-life, as translation efficiency has shown to influence mRNA degradation in *E. coli* (Fry *et al.*, 1972, Pato *et al*., 1973). We therefore compared base pairing in the ribosome binding site (RBS) of leadered transcripts to mRNA decay rates and ribosome occupancy.

RBS analysis was done in *M. smegmatis*, due to the availability of ribosome occupancy data that were not available for MTB. To analyze degradation rates, *M. smegmatis* transcripts were clustered into slow (S), medium (M), and fast (F) decay groups based on degradation profiles by Huaming Sun of the Shell Lab. The clustered data was chosen for analysis here because previously reported data in this project regarding unpaired start codons of leaderless transcripts was shown to be an influential factor on characterizing slow versus fast decay clusters by Huaming. In addition, ribosome profiling data were obtained from published work (Chen *et al.,* 2020), and occupancy at the last 20 nucleotides of the 5' UTR plus the first 18 nucleotides of the coding sequence were determined within the Shell lab.

To assess the effect of translation efficiency on mRNA stability, pairing probabilities and ribosome occupancy were compared to patterns of decay. To predict pairing at the RBS, the last 30 nucleotides of the 5' UTR plus the first 20 nucleotides of the CDS or the start codon were folded. Probabilities of unpaired regions for each transcript were calculated by taking the average probability that each nucleotide and the previous nucleotide are both unpaired. This was done across the whole region that was allowed to fold, as well as within the Shine Dalgarno sequence, represented by the -6 to -14 region relative to the start codon. The probability that the whole start codon was unpaired was also assessed.

We hypothesized that unpaired RBS regions are more accessible to ribosomes, and that ribosome binding protects the transcripts from degradation. In support of this hypothesis, results show that transcripts with slower decay rates are more likely to have start codons predicted to be unpaired. This was true when allowing the last 30 nucleotides of the 5' UTR plus either the start codon (Figure 14 A) or the first 20 nucleotides of the coding sequence (Figure 15 A) to fold.

Further supporting this hypothesis, when the last 30 nucleotides of the 5' UTR plus the first 20 nucleotides of the coding sequence (Figure 15 E) or when the first 20 nucleotides of leaderless transcripts (Figure 15 G) were allowed to fold, there is a significant, albeit weak, positive correlation between probability of unpaired start codons and ribosome occupancy at the -20 to +18 region relative to the start codon. Similar correlations were observed between the average unpaired probability of the whole region allowed to fold and ribosome occupancy at the same region when the last 30 nucleotides of the 5' UTR plus either the start codon (Figure 14 F) or the first 20 nucleotides (Figure 15 C) of the coding sequence were allowed to fold. This supports the hypothesis that unpaired RBSs allow ribosomes better access to the transcript and protect them from degradation.

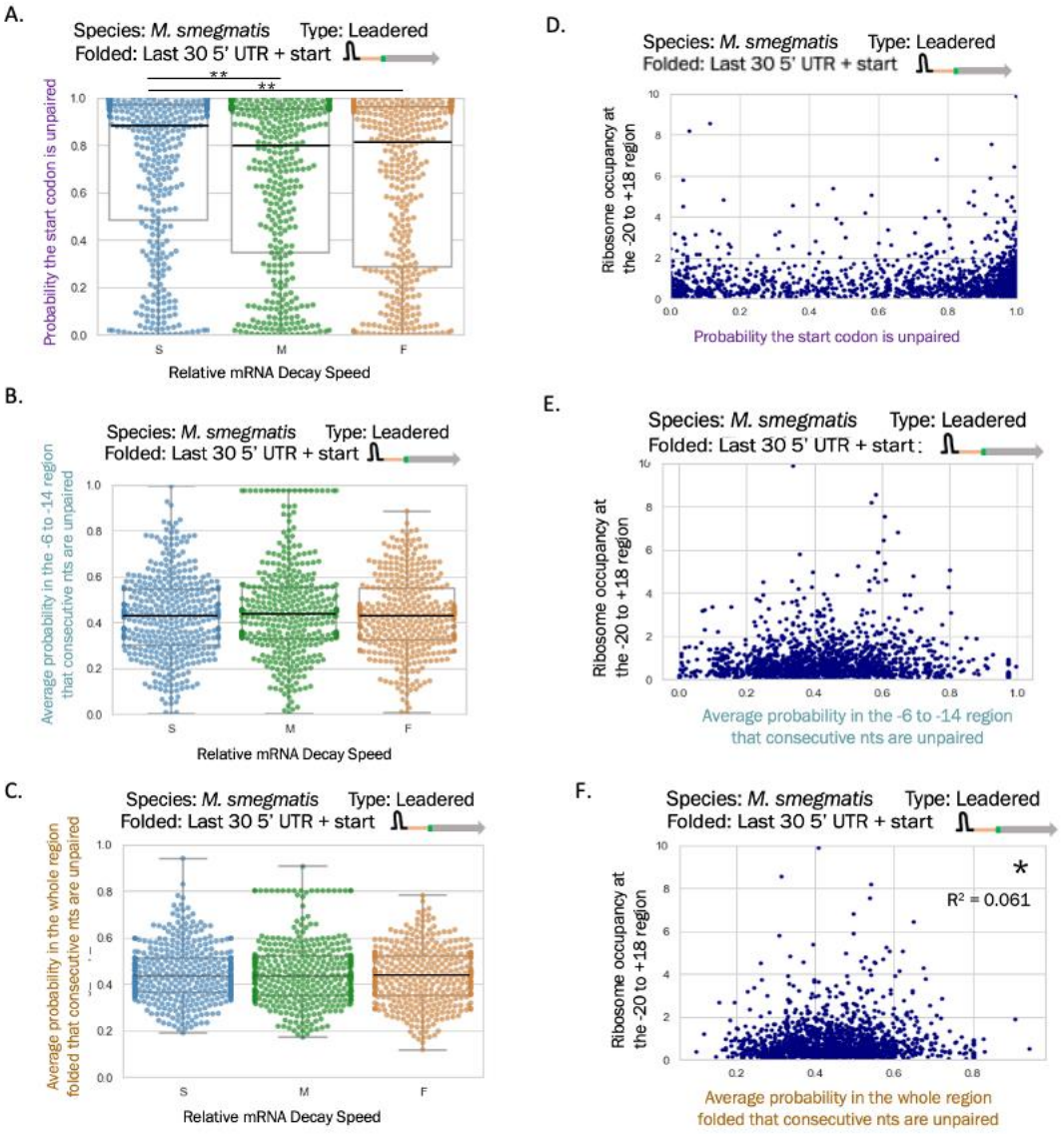


Figure 14. Base pairing within the last 30 nucleotides of the 5' UTR plus the start codon was compared to decay pattern and ribosome occupancy in leadered transcripts.
Local MFE secondary structures were predicted within the last 30 nucleotides of the 5' UTR plus the start codon of leadered *M. smegmatis* transcripts using ViennaRNA. (A-C) Probabilities that the (A) start codon and the average probability of having two consecutive unpaired nucleotides within (B) the -6 to -14 region or (C) the whole region folded were compared for transcripts having slow (S), medium (M), and Fast (F) mRNA decay using Kruskal-Wallis analysis of variance followed by Dunn's post-hoc analysis. (D-F) Ribosome occupancy at the -20 to +18 region of the mRNA was assessed compared to unpaired probabilities of (D) the start codon, (E) the -6 to -14 region, and (F) the whole region, as described previously, using Spearman's correlation analysis. *, $P < 0.05$; **, $P < 0.01$.

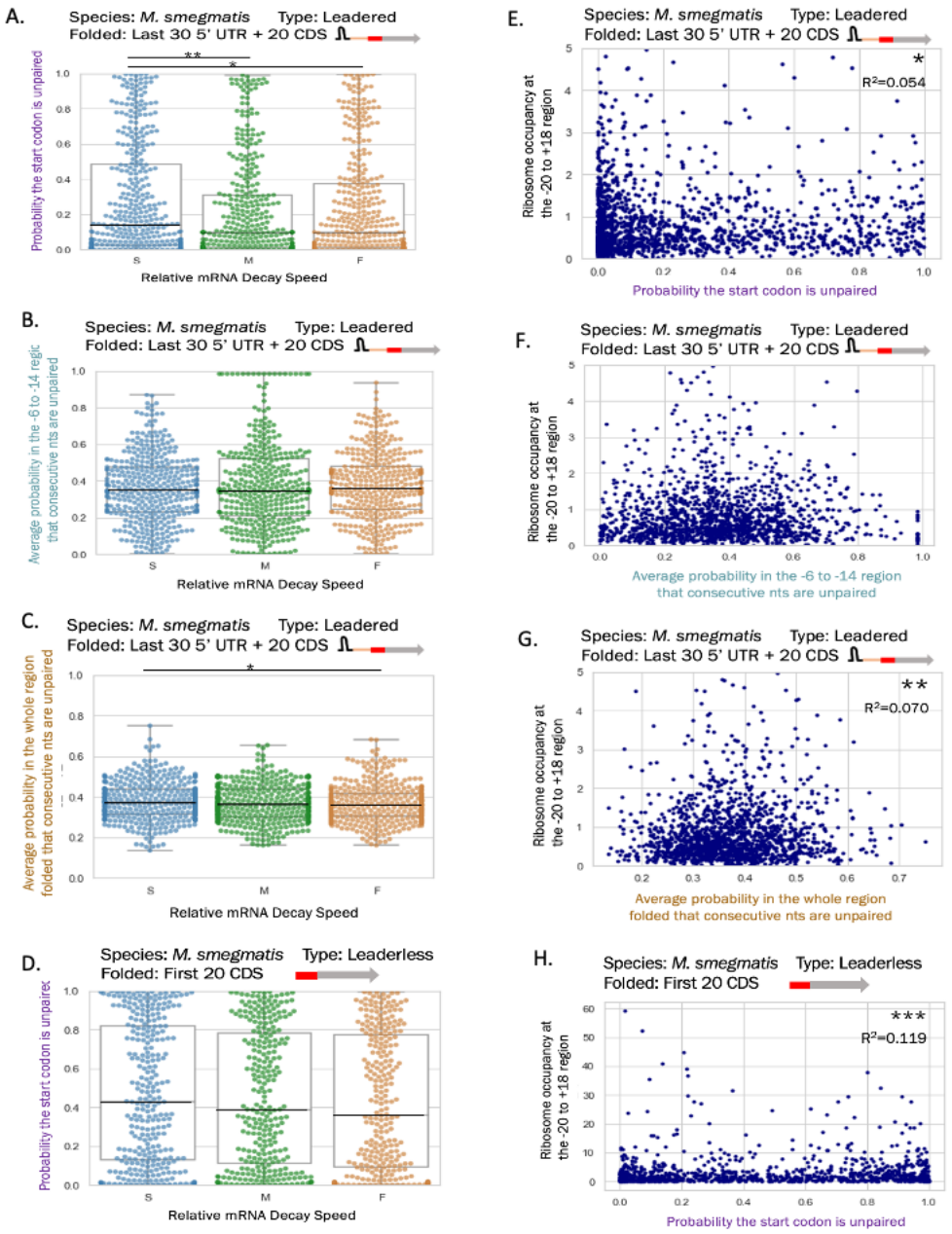


Figure 15. Base pairing within the last 30 nucleotides of the 5' UTR (leadered) plus the first 20 nucleotides of the coding sequence (leadered and leaderless) was compared to decay pattern and ribosome occupancy. Local MFE secondary structures were predicted within the last 30 nucleotides of the 5' UTR plus the first 20 of the CDS (leadered) or just the first 20 of the CDS (leaderless) in *M. smegmatis* transcripts using ViennaRNA. (A-D) Probabilities that the (A&D) start codon and the average probability of having two consecutive unpaired nucleotides within (B) the -6 to -14 region or (C) the whole region folded were compared for transcripts having slow (S), medium (M), and Fast (F) mRNA decay using Kruskal-Wallis analysis of variance followed by Dunn's post-hoc analysis. (E-H) Ribosome occupancy at the -20 to +18 region of the mRNA was also assessed compared to unpaired probabilities of (E&H) the start codon, (F) the -6 to -14 region, and (G) the whole region, as described previously, using Spearman's correlation analysis. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

Discussion

Here we investigated the influence of mRNA secondary structure on degradation rate in mycobacteria. The results of this project suggest that the hairpin loops at the 5' end of the *esxB* UTR are not a significant factor in dictating the half-life of the transcript. While deletion of either hairpin loop did not have a significant effect on *yfp* expression compared to the full 5' UTR, deletion of both loops surprisingly increased *yfp* RNA and protein expression. In contrast, there was similarity in half-life between transcripts with the native *esxB* 5' UTR and deletion of either or both 5' loops, leading us to suggest the increase in *yfp* expression is caused by increased transcription of *yfp* when both 5' hairpin loops of the *esxB* 5' UTR are deleted. One possible explanation for this effect is the loops acting as transcriptional terminators. Overall, these results show that the 5' hairpin loops of *esxB* do not confer protection from mRNA degradation, as we initially hypothesized.

Analysis of the *M. smegmatis* and MTB transcriptomes, similarly, did not reveal significant relationships between unpaired nucleotides at the beginning of 5' UTRs and mRNA stability. However, the relationship between unpaired start codons of leaderless transcripts and mRNA decay rate approached significance and was found to be one of the ten most influential factors in predicting decay from other work within the Shell lab. Further analysis of base pairing probabilities at ribosome binding sites in leadered and leaderless transcripts in *M. smegmatis* revealed significant relationship between an unpaired start codons and slow decay rates. This suggests that when ribosomes have easier access to the Shine Dalgarno sequence in leadered transcripts and start codons in leaderless transcripts, the ribosomes associate with the transcripts and protect them from degradation. This is supported by analysis of ribosome occupancy, which showed a positive correlation with unpaired regions at the RBS. Furthermore, this is consistent with studies in *E. coli* in which translation inhibitors that allow ribosomes to stall on transcripts resulted in increased mRNA stability (Fry *et al*., 1972, Pato *et al*., 1973). Taken together, these results imply that secondary structures in the RBS have a greater influence on mRNA decay than structures at the 5' end of transcripts in mycobacteria.

These results differ from previous work suggesting that RNase E requires at least five consecutive unpaired nucleotides at the 5' end (Emory *et al*., 1992, Bouvet & Belasco, 1992) and that pyrophosphate removal by RppH requires an unpaired 5' terminus (Celesnik *et al*., 2007) for efficient transcript degradation in *E. coli*. Therefore, we suggest that mRNA degradation mechanisms in mycobacteria are differently influenced by secondary structures, possibly due to the RNase E and/or as-yet unidentified RppH ortholog being able to engage the 5' end regardless of 5' structures. Given the complex nature of transcriptional regulation, there are many more factors involved in mRNA degradation other than secondary structure near the 5' terminus. In a genome-wide study on determinants of mRNA degradation in *E. coli*, Esquerre *et al.* reported that mRNA concentration was the factor most predictive of half-life, with GC content, 5' UTR sequence motifs and structure, gene function, and essentiality amongst other factors found to be predictive in a significant way (Esquerre *et al.*, 2015). From this study, it is clear that the influence of 5' secondary structures on mRNA degradation is different in mycobacteria than in *E. coli.* Future studies will need to be done to determine which other factors may be more influential on mycobacterial mRNA degradation mechanisms.

References

Amilon, K.R., Letley, D.P., Winter, J.A., Robinson, K., and Atherton, J.C. (2015). Expression of the Helicobacter pylori virulence factor vacuolating cytotoxin A (vacA) is influenced by a potential stem-loop structure in the 5' untranslated region of the transcript. *Molecular Microbiology*, 98(5), 831-846. doi: 10.1111/mmi.13160.

Bouvet, P., & Belasco, J. G. (1992). Control of RNase E-mediated RNA degradation by 5′-terminal base pairing in E. coil. *Nature*, 360(6403), 488–491. https://doi.org/10.1038/360488a0

Braun, F., Le Derout, J., and Regnier, P. (1998). Ribosomes inhibit an RNase E cleavage which induces the decay of the rpsO mRNA of Escherichia coli. *EMBO J* 17(16)**,** 4790-4797. doi: 10.1093/emboj/17.16.4790.

Carrier, T. A., & Keasling, J. D. (1999). Library of synthetic 5' secondary structures to manipulate mRNA stability in Escherichia coli. *Biotechnology progress*, 15(1), 58–64. https://doi.org/10.1021/bp9801143

Celesnik, H., Deana, A., & Belasco, J. G. (2007). Initiation of RNA decay in Escherichia coli by 5' pyrophosphate removal. *Molecular cell*, 27(1), 79–90. https://doi.org/10.1016/j.molcel.2007.05.038

Chen, Y. X., Xu, Z. Y., Ge, X., Sanyal, S., Lu, Z. J., & Javid, B. (2020). Selective translation by alternative bacterial ribosomes. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(32), 19487–19496. https://doi.org/10.1073/pnas.2009607117

Ching, C., Gozzi, K., Heinemann, B., Chai, Y., and Godoy, V.G. (2017). RNA-Mediated cis Regulation in Acinetobacter baumannii Modulates Stress-Induced Phenotypic Variation. *J Bacteriol* 199(11). doi: 10.1128/JB.00799-16.

Cohen,S.N. and McDowall,K.J. (1997) RNase E: still a wonderfully mysterious enzyme. *Molecular Microbiology*, 23, 1099–1106.

Condon, C. (2003). RNA Processing and Degradation in Bacillus subtilis. *Microbiol Mol Biol Rev*, 62(2), 157-174. doi:10.1128/MMBR.67.2.157-174.2003

Czyz, A., Mooney, R. A., Iaconi, A., & Landick, R. (2014). Mycobacterial RNA Polymerase Requires a U-Tract at Intrinsic Terminators and Is Aided by NusG at Suboptimal Terminators. *MBio*, 5(2), e00931-14. https://doi.org/10.1128/mBio.00931-14

Deana, A., Celesnik, H., & Belasco, J. G. (2008). The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature*, 451(7176), 355–358. https://doi.org/10.1038/nature06475

de Rivera, J. (2016). The Effects of Post-Transcriptional Processing on mRNA Stability in M. smegmatis: Worcester Polytechnic Institute.

Ehrt, S., Guo, X. V., Hickey, C. M., Ryou, M., Monteleone, M., Riley, L. W., & Schnappinger, D. (2005). Controlling gene expression in mycobacteria with anhydrotetracycline and Tet repressor. *Nucleic Acids Research*, 33(2), e21–e21. https://doi.org/10.1093/nar/gni013

Emory, S. A., Bouvet, P., & Belasco, J. G. (1992). A 5'-terminal stem-loop structure can stabilize mRNA in *Escherichia coli*. *Genes & Development*, 6(1), 135–148. https://doi.org/10.1101/gad.6.1.135

Esquerré, T., Moisan, A., Chiapello, H. et al. Genome-wide investigation of mRNA lifetime determinants in Escherichia coli cells cultured at different growth rates. *BMC Genomics* 16, 275 (2015). https://doi.org/10.1186/s12864-015-1482-8

Global tuberculosis report 2020. (2020) Geneva: World Health Organization; 2020 Licence: CC BY-NC-SA 3.0 IGO.

Hambraeus, G., Persson, M., & Rutberg, B. (2000). The aprE leader is a determinant of extreme mRNA stability in Bacillus subtilis. *Microbiology*, 146(12), 3051–3059. https://doi.org/10.1099/00221287-146-12-3051

Huff, J., Czyz, A., Landick, R., & Niederweis, M. (2010). Taking phage integration to the next level as a genetic tool for mycobacteria. *Gene*, 468(1), 8–19. https://doi.org/10.1016/j.gene.2010.07.012

Klug, G., and Cohen, S.N. (1990). Combined actions of multiple hairpin loop structures and sites of rate-limiting endonucleolytic cleavage determine differential degradation rates of individual segments within polycistronic puf operon mRNA. *J Bacteriol* 172(9), 5140- 5146. doi: 10.1128/jb.172.9.5140-5146.1990.

Kristiansen, K. I., Weel-Sneve, R., Booth, J. A., & Bjørås, M. (2016). Mutually exclusive RNA secondary structures regulate translation initiation of DinQ in Escherichia coli. *RNA* (New York, N.Y.), 22(11), 1739–1749. https://doi.org/10.1261/rna.058461.116

Kovacs, L., Csanadi, A., Megyeri, K., Kaberdin, V.R. and Miczak,A. (2005) Mycobacterial RNase E-associated proteins. *Microbiol. Immunol*., 49, 1003–1007.

Lorenz, R., Bernhart, S.H., Höner zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F. and Hofacker, I.L. (2011). ViennaRNA Package 2.0. *Algorithms for Molecular Biology*, 6:1. doi:10.1186/1748-7188-6-26

Ma, Y., Keil, V., & Sun, J. (2015). Characterization of Mycobacterium tuberculosis EsxA membrane insertion: roles of N- and C-terminal flexible arms and central helix-turn-helix motif. *The Journal of biological chemistry*, *290*(11), 7314–7322. https://doi.org/10.1074/jbc.M114.622076

Martini, M. C., Zhou, Y., Sun, H., & Shell, S. S. (2019). Defining the Transcriptional and Post-transcriptional Landscapes of Mycobacterium smegmatis in Aerobic Growth and Hypoxia. *Frontiers in microbiology*, 10, 591. https://doi.org/10.3389/fmicb.2019.00591

Morris, P., Marinelli, L. J., Jacobs-Sera, D., Hendrix, R. W., & Hatfull, G. F. (2008). Genomic Characterization of Mycobacteriophage Giles: Evidence for Phage Acquisition of Host DNA by Illegitimate Recombination. *Journal of Bacteriology*, 190(6), 2172–2182. https://doi.org/10.1128/JB.01657-07

Nahvi, A., Sudarsan, N., Ebert, M. S., Zou, X., Brown, K. L., & Breaker, R. R. (2002). Genetic control by a metabolite binding mRNA. *Chemistry & biology*, 9(9), 1043. https://doi.org/10.1016/s1074-5521(02)00224-7

Nguyen, T. (2019). Investigating the Post-Transcriptional Effects of the sigA 5' UTR on Gene Expression: Worcester Polytechnic Institute.

Nguyen, T.G., Vargas-Blanco, D.A., Roberts, L.A., and Shell, S.S. (2020). The impact of leadered and leaderless gene structures on translation efficiency, transcript stability, and predicted transcription rates in Mycobacterium smegmatis. *J Bacteriol*. https://doi.org/10.1128/JB.00746-19.

Obana, N., Shirahama, Y., Abe, K., & Nakamura, K. (2010). Stabilization of Clostridium perfringens collagenase mRNA by VR-RNA-dependent cleavage in 5′ leader sequence. Molecular Microbiology, 77(6), 1416–1428. https://doi.org/10.1111/j.1365-2958.2010.07258.x

Pato, M.L., Bennett, P.M., and von Meyenburg, K. (1973). Messenger ribonucleic acid synthesis and degradation in Escherichia coli during inhibition of translation. *J Bacteriol* 116(2), 710-718.

Pym, A. S., Brodin, P., Majlessi, L., Brosch, R., Demangel, C., Williams, A., Griffiths, K. E., Marchal, G., Leclerc, C., & Cole, S. T. (2003). Recombinant BCG exporting ESAT-6 confers enhanced protection against tuberculosis. *Nature Medicine*, 9(5), 533–539. https://doi.org/10.1038/nm859

Rustad, T. R., Minch, K. J., Brabant, W., Winkler, J. K., Reiss, D. J., Baliga, N. S., & Sherman, D. R. (2013). Global analysis of mRNA stability in Mycobacterium tuberculosis. *Nucleic Acids Research*, 41(1), 509–517. https://doi.org/10.1093/nar/gks1019

Sala, C., Forti, F., Magnoni, F., & Ghisotti, D. (2008). The katG mRNA of Mycobacterium tuberculosis and Mycobacterium smegmatis is processed at its 5' end and is stabilized by both a polypurine sequence and translation initiation. *BMC molecular biology*, 9, 33. https://doi.org/10.1186/1471-2199-9-33

Schneider, E., Blundell, M., and Kennell, D. (1978). Translation and mRNA decay. *Mol Gen Genet* 160(2), 121-129.

Shell, S. S., Chase, M. R., Gray, T. A., Wade, J. T., Singh, N., DeJesus, M., Sacchettini, J. C., Ioerger, T. R., & Fortune, S. M. Submitted Manuscript. mRNA cleavage shapes mycobacterial transcriptomes and stabilizes the esxB-esxA virulence factor transcript.

Shell, S. S., Wang, J., Lapierre, P., Mir, M., Chase, M. R., Pyle, M. M., Gawande, R., Ahmad, R., Sarracino, D. A., Ioerger, T. R., Fortune, S. M., Derbyshire, K. M., Wade, J. T., & Gray, T. A. (2015). Leaderless Transcripts and Small Proteins Are Common Features of the Mycobacterial Translational Landscape. *PLoS genetics*, 11(11), e1005641. https://doi.org/10.1371/journal.pgen.1005641

Taverniti, V., Forti, F., Ghisotti, D., & Putzer, H. (2011). Mycobacterium smegmatis RNase J is a 5'-3' exo-/endoribonuclease and both RNase J and RNase E are involved in ribosomal RNA maturation. *Molecular microbiology*, *82*(5), 1260–1276. https://doi.org/10.1111/j.1365-2958.2011.07888.x

Appendix A.

Leadered Genes Used				Leaderless Genes Used	
Gene - *M. smegmatis*	Size of 5'UTR	Gene - MTB	Size of 5'UTR	Gene - *M. smegmatis*	Gene - MTB
MSMEG_0001	105	Rv0001	263	MSMEG_0024	Rv0013
MSMEG_0005	57	Rv0002	117	MSMEG_0029	Rv0023
MSMEG_0007	287	Rv0005	57	MSMEG_0042	Rv0029
MSMEG_0015	26	Rv0008c	56	MSMEG_0073	Rv0030
MSMEG_0016	80	Rv0011c	90	MSMEG_0074	Rv0036c
MSMEG_0020	103	Rv0012	22	MSMEG_0092	Rv0038
MSMEG_0021	65	Rv0014c	64	MSMEG_0098	Rv0041
MSMEG_0023	32	Rv0015c	270	MSMEG_0104	Rv0042c
MSMEG_0026	54	Rv0017c	117	MSMEG_0111	Rv0044c
MSMEG_0027	48	Rv0018c	190	MSMEG_0119	Rv0048c
MSMEG_0028	213	Rv0019c	59	MSMEG_0124	Rv0049
MSMEG_0030	302	Rv0020c	71	MSMEG_0127	Rv0062
MSMEG_0032	253	Rv0021c	40	MSMEG_0128	Rv0072
MSMEG_0033	123	Rv0025	124	MSMEG_0129	Rv0076c
MSMEG_0034	54	Rv0026	37	MSMEG_0146	Rv0078A
MSMEG_0036	294	Rv0027	111	MSMEG_0148	Rv0093c
MSMEG_0044	95	Rv0032	355	MSMEG_0180	Rv0102
MSMEG_0046	66	Rv0037c	29	MSMEG_0188	Rv0118c
MSMEG_0047	54	Rv0043c	41	MSMEG_0212	Rv0128
MSMEG_0048	23	Rv0070c	417	MSMEG_0213	Rv0133
MSMEG_0049	402	Rv0075	72	MSMEG_0215	Rv0140
MSMEG_0050	337	Rv0078	66	MSMEG_0224	Rv0141c
MSMEG_0051	43	Rv0084	122	MSMEG_0229	Rv0142
MSMEG_0052	61	Rv0091	62	MSMEG_0230	Rv0143c
MSMEG_0053	43	Rv0094c	296	MSMEG_0231	Rv0144
MSMEG_0055	19	Rv0096	56	MSMEG_0234	Rv0154c
MSMEG_0056	71	Rv0098	309	MSMEG_0235	Rv0161
MSMEG_0060	107	Rv0099	9	MSMEG_0239	Rv0163
MSMEG_0061	11	Rv0103c	26	MSMEG_0248	Rv0164
MSMEG_0067	28	Rv0107c	26	MSMEG_0252	Rv0177
MSMEG_0069	75	Rv0109	42	MSMEG_0257	Rv0184
MSMEG_0070	462	Rv0110	264	MSMEG_0261	Rv0187
MSMEG_0076	120	Rv0111	105	MSMEG_0287	Rv0189c
MSMEG_0078	182	Rv0120c	54	MSMEG_0289	Rv0199
MSMEG_0079	246	Rv0129c	53	MSMEG_0301	Rv0204c

MSMEG_0083	79	Rv0130	17	MSMEG_0310	Rv0205
MSMEG_0087	42	Rv0131c	12	MSMEG_0319	Rv0213c
MSMEG_0093	1	Rv0134	3	MSMEG_0356	Rv0214
MSMEG_0095	42	Rv0136	58	MSMEG_0360	Rv0217c
MSMEG_0096	41	Rv0137c	11	MSMEG_0361	Rv0220
MSMEG_0101	23	Rv0138	27	MSMEG_0364	Rv0224c
MSMEG_0107	4	Rv0139	45	MSMEG_0367	Rv0228
MSMEG_0109	228	Rv0147	7	MSMEG_0373	Rv0230c
MSMEG_0110	27	Rv0152c	126	MSMEG_0391	Rv0237
MSMEG_0117	29	Rv0155	41	MSMEG_0393	Rv0258c
MSMEG_0118	19	Rv0156	492	MSMEG_0394	Rv0262c
MSMEG_0120	3	Rv0158	356	MSMEG_0407	Rv0264c
MSMEG_0130	75	Rv0175	165	MSMEG_0423	Rv0265c
MSMEG_0131	76	Rv0176	282	MSMEG_0434	Rv0269c
MSMEG_0132	135	Rv0178	179	MSMEG_0436	Rv0270
MSMEG_0133	62	Rv0179c	5	MSMEG_0437	Rv0277c
MSMEG_0140	101	Rv0182c	439	MSMEG_0438	Rv0281
MSMEG_0142	111	Rv0183	1	MSMEG_0441	Rv0295c
MSMEG_0143	283	Rv0186	166	MSMEG_0448	Rv0307c
MSMEG_0145	258	Rv0193c	51	MSMEG_0464	Rv0310c
MSMEG_0152	84	Rv0202c	43	MSMEG_0469	Rv0311
MSMEG_0153	84	Rv0206c	207	MSMEG_0471	Rv0315
MSMEG_0155	36	Rv0209	18	MSMEG_0491	Rv0321
MSMEG_0166	28	Rv0210	108	MSMEG_0492	Rv0332
MSMEG_0167	121	Rv0211	106	MSMEG_0538	Rv0333
MSMEG_0168	95	Rv0215c	28	MSMEG_0541	Rv0359
MSMEG_0173	360	Rv0216	10	MSMEG_0544	Rv0364
MSMEG_0174	367	Rv0225	1	MSMEG_0546	Rv0377
MSMEG_0179	29	Rv0226c	240	MSMEG_0548	Rv0379
MSMEG_0181	29	Rv0227c	191	MSMEG_0551	Rv0390
MSMEG_0186	46	Rv0231	33	MSMEG_0561	Rv0394c
MSMEG_0191	100	Rv0232	23	MSMEG_0563	Rv0400c
MSMEG_0194	40	Rv0234c	9	MSMEG_0564	Rv0401
MSMEG_0195	456	Rv0235c	27	MSMEG_0578	Rv0406c
MSMEG_0203	107	Rv0236A	40	MSMEG_0580	Rv0407
MSMEG_0206	11	Rv0239	17	MSMEG_0592	Rv0413
MSMEG_0208	40	Rv0245	39	MSMEG_0598	Rv0427c
MSMEG_0210	358	Rv0252	13	MSMEG_0599	Rv0429c
MSMEG_0218	30	Rv0256c	292	MSMEG_0607	Rv0432

MSMEG_0222	38	Rv0266c	196	MSMEG_0609	Rv0439c
MSMEG_0226	140	Rv0268c	26	MSMEG_0612	Rv0443
MSMEG_0227	81	Rv0271c	38	MSMEG_0614	Rv0452
MSMEG_0228	11	Rv0275c	35	MSMEG_0625	Rv0455c
MSMEG_0232	3	Rv0293c	10	MSMEG_0629	Rv0457c
MSMEG_0233	22	Rv0305c	35	MSMEG_0632	Rv0460
MSMEG_0240	175	Rv0306	156	MSMEG_0633	Rv0462
MSMEG_0241	69	Rv0312	34	MSMEG_0634	Rv0465c
MSMEG_0242	214	Rv0338c	46	MSMEG_0635	Rv0471c
MSMEG_0243	83	Rv0339c	17	MSMEG_0636	Rv0479c
MSMEG_0245	142	Rv0343	335	MSMEG_0645	Rv0480c
MSMEG_0250	457	Rv0348	39	MSMEG_0668	Rv0481c
MSMEG_0253	97	Rv0349	107	MSMEG_0678	Rv0483
MSMEG_0254	10	Rv0355c	83	MSMEG_0682	Rv0485
MSMEG_0255	65	Rv0365c	93	MSMEG_0683	Rv0486
MSMEG_0264	186	Rv0369c	36	MSMEG_0694	Rv0489
MSMEG_0268	27	Rv0383c	1	MSMEG_0702	Rv0495c
MSMEG_0270	116	Rv0388c	244	MSMEG_0717	Rv0496
MSMEG_0282	68	Rv0392c	30	MSMEG_0719	Rv0500
MSMEG_0285	54	Rv0398c	62	MSMEG_0721	Rv0503c
MSMEG_0290	152	Rv0402c	143	MSMEG_0724	Rv0524
MSMEG_0296	21	Rv0403c	126	MSMEG_0731	Rv0539
MSMEG_0299	65	Rv0404	223	MSMEG_0742	Rv0546c
MSMEG_0300	126	Rv0412c	44	MSMEG_0750	Rv0547c
MSMEG_0303	16	Rv0415	97	MSMEG_0754	Rv0550c
MSMEG_0305	34	Rv0418	1	MSMEG_0756	Rv0551c
MSMEG_0309	35	Rv0419	35	MSMEG_0762	Rv0554
MSMEG_0311	1	Rv0420c	69	MSMEG_0765	Rv0559c
MSMEG_0314	84	Rv0423c	97	MSMEG_0768	Rv0561c
MSMEG_0317	58	Rv0425c	42	MSMEG_0775	Rv0562
MSMEG_0321	44	Rv0426c	114	MSMEG_0789	Rv0576
MSMEG_0322	63	Rv0428c	386	MSMEG_0790	Rv0580c
MSMEG_0331	28	Rv0430	73	MSMEG_0810	Rv0581
MSMEG_0341	124	Rv0442c	92	MSMEG_0812	Rv0586
MSMEG_0343	2	Rv0445c	1	MSMEG_0816	Rv0599c
MSMEG_0348	146	Rv0446c	265	MSMEG_0825	Rv0605
MSMEG_0358	29	Rv0449c	397	MSMEG_0832	Rv0618
MSMEG_0359	178	Rv0450c	156	MSMEG_0835	Rv0623
MSMEG_0363	8	Rv0451c	184	MSMEG_0838	Rv0626

MSMEG_0366	41	Rv0456A	124	MSMEG_0854	Rv0628c
MSMEG_0368	32	Rv0464c	232	MSMEG_0857	Rv0632c
MSMEG_0372	34	Rv0466	53	MSMEG_0877	Rv0633c
MSMEG_0376	37	Rv0470c	76	MSMEG_0884	Rv0634A
MSMEG_0380	98	Rv0472c	25	MSMEG_0886	Rv0634c
MSMEG_0381	68	Rv0476	286	MSMEG_0887	Rv0635
MSMEG_0382	434	Rv0484c	478	MSMEG_0889	Rv0647c
MSMEG_0383	122	Rv0488	495	MSMEG_0890	Rv0660c
MSMEG_0384	84	Rv0490	59	MSMEG_0903	Rv0661c
MSMEG_0386	3	Rv0497	276	MSMEG_0906	Rv0670
MSMEG_0387	139	Rv0500A	70	MSMEG_0915	Rv0672
MSMEG_0388	116	Rv0500B	28	MSMEG_0917	Rv0677c
MSMEG_0392	150	Rv0502	77	MSMEG_0920	Rv0678
MSMEG_0398	135	Rv0504c	147	MSMEG_0923	Rv0688
MSMEG_0399	81	Rv0505c	46	MSMEG_0928	Rv0695
MSMEG_0402	307	Rv0506	97	MSMEG_0929	Rv0698
MSMEG_0404	14	Rv0509	44	MSMEG_0930	Rv0712
MSMEG_0406	84	Rv0520	51	MSMEG_0933	Rv0713
MSMEG_0408	79	Rv0522	130	MSMEG_0935	Rv0729
MSMEG_0409	59	Rv0523c	93	MSMEG_0936	Rv0744c
MSMEG_0410	492	Rv0529	140	MSMEG_0938	Rv0748
MSMEG_0411	61	Rv0530	74	MSMEG_0939	Rv0772
MSMEG_0412	28	Rv0533c	71	MSMEG_0943	Rv0774c
MSMEG_0413	3	Rv0537c	242	MSMEG_0949	Rv0777
MSMEG_0415	3	Rv0538	105	MSMEG_0964	Rv0780
MSMEG_0416	32	Rv0543c	49	MSMEG_0977	Rv0784
MSMEG_0419	307	Rv0545c	75	MSMEG_0982	Rv0786c
MSMEG_0420	26	Rv0548c	23	MSMEG_0988	Rv0792c
MSMEG_0422	38	Rv0553	90	MSMEG_1010	Rv0799c
MSMEG_0428	233	Rv0557	379	MSMEG_1029	Rv0801
MSMEG_0429	91	Rv0563	42	MSMEG_1036	Rv0802c
MSMEG_0433	282	Rv0571c	94	MSMEG_1038	Rv0803
MSMEG_0443	23	Rv0572c	29	MSMEG_1039	Rv0811c
MSMEG_0447	43	Rv0584	333	MSMEG_1042	Rv0817c
MSMEG_0449	37	Rv0585c	56	MSMEG_1046	Rv0827c
MSMEG_0450	317	Rv0588	27	MSMEG_1049	Rv0831c
MSMEG_0451	31	Rv0603	40	MSMEG_1061	Rv0838
MSMEG_0453	29	Rv0610c	334	MSMEG_1067	Rv0844c
MSMEG_0455	132	Rv0612	249	MSMEG_1068	Rv0853c

MSMEG_0457	34	Rv0615	9	MSMEG_1073	Rv0854
MSMEG_0462	50	Rv0621	12	MSMEG_1075	Rv0858c
MSMEG_0465	25	Rv0638	61	MSMEG_1077	Rv0861c
MSMEG_0466	32	Rv0639	150	MSMEG_1098	Rv0880
MSMEG_0473	166	Rv0642c	53	MSMEG_1103	Rv0884c
MSMEG_0480	15	Rv0643c	42	MSMEG_1115	Rv0893c
MSMEG_0484	294	Rv0644c	83	MSMEG_1117	Rv0898c
MSMEG_0493	38	Rv0645c	15	MSMEG_1123	Rv0903c
MSMEG_0494	37	Rv0651	216	MSMEG_1124	Rv0904c
MSMEG_0499	52	Rv0669c	55	MSMEG_1131	Rv0905
MSMEG_0500	81	Rv0675	282	MSMEG_1132	Rv0907
MSMEG_0504	62	Rv0680c	55	MSMEG_1133	Rv0919
MSMEG_0519	15	Rv0681	134	MSMEG_1136	Rv0921
MSMEG_0520	56	Rv0682	183	MSMEG_1140	Rv0923c
MSMEG_0530	77	Rv0684	165	MSMEG_1173	Rv0933
MSMEG_0531	212	Rv0685	73	MSMEG_1175	Rv0937c
MSMEG_0537	194	Rv0686	29	MSMEG_1194	Rv0938
MSMEG_0539	36	Rv0687	16	MSMEG_1210	Rv0946c
MSMEG_0540	19	Rv0690c	199	MSMEG_1213	Rv0948c
MSMEG_0545	51	Rv0691c	30	MSMEG_1222	Rv0956
MSMEG_0547	8	Rv0692	118	MSMEG_1225	Rv0960
MSMEG_0549	223	Rv0700	284	MSMEG_1226	Rv0961
MSMEG_0557	137	Rv0702	92	MSMEG_1250	Rv0967
MSMEG_0559	44	Rv0711	34	MSMEG_1261	Rv0969
MSMEG_0560	56	Rv0714	145	MSMEG_1265	Rv0981
MSMEG_0566	65	Rv0720	320	MSMEG_1277	Rv0989c
MSMEG_0572	64	Rv0725c	66	MSMEG_1279	Rv0992c
MSMEG_0574	390	Rv0726c	34	MSMEG_1283	Rv0997
MSMEG_0575	104	Rv0737	497	MSMEG_1293	Rv1000c
MSMEG_0581	25	Rv0740	237	MSMEG_1305	Rv1001
MSMEG_0583	111	Rv0741	109	MSMEG_1331	Rv1005c
MSMEG_0586	115	Rv0743c	260	MSMEG_1332	Rv1011
MSMEG_0595	19	Rv0749A	42	MSMEG_1334	Rv1019
MSMEG_0596	68	Rv0750	4	MSMEG_1353	Rv1032c
MSMEG_0600	96	Rv0756c	271	MSMEG_1363	Rv1055
MSMEG_0602	164	Rv0757	60	MSMEG_1380	Rv1059
MSMEG_0603	46	Rv0761c	66	MSMEG_1383	Rv1063c
MSMEG_0605	242	Rv0762c	12	MSMEG_1385	Rv1075c
MSMEG_0606	43	Rv0766c	102	MSMEG_1387	Rv1077

MSMEG_0608	47	Rv0775	126	MSMEG_1391	Rv1086
MSMEG_0615	114	Rv0781	49	MSMEG_1414	Rv1092c
MSMEG_0623	205	Rv0783c	3	MSMEG_1416	Rv1093
MSMEG_0638	3	Rv0785	90	MSMEG_1418	Rv1094
MSMEG_0641	150	Rv0787A	70	MSMEG_1425	Rv1096
MSMEG_0643	33	Rv0794c	26	MSMEG_1432	Rv1099c
MSMEG_0656	6	Rv0805	39	MSMEG_1476	Rv1100
MSMEG_0660	31	Rv0807	3	MSMEG_1492	Rv1102c
MSMEG_0661	75	Rv0808	81	MSMEG_1512	Rv1104
MSMEG_0662	193	Rv0809	1	MSMEG_1526	Rv1109c
MSMEG_0663	40	Rv0823c	279	MSMEG_1529	Rv1111c
MSMEG_0667	80	Rv0826	59	MSMEG_1561	Rv1112
MSMEG_0671	39	Rv0832	77	MSMEG_1565	Rv1118c
MSMEG_0673	93	Rv0834c	41	MSMEG_1567	Rv1130
MSMEG_0674	42	Rv0835	161	MSMEG_1592	Rv1132
MSMEG_0676	23	Rv0837c	44	MSMEG_1601	Rv1142c
MSMEG_0679	176	Rv0839	26	MSMEG_1610	Rv1151c
MSMEG_0689	72	Rv0852	9	MSMEG_1621	Rv1152
MSMEG_0690	75	Rv0871	172	MSMEG_1623	Rv1155
MSMEG_0692	28	Rv0879c	3	MSMEG_1631	Rv1159A
MSMEG_0698	38	Rv0890c	422	MSMEG_1635	Rv1160
MSMEG_0700	490	Rv0891c	284	MSMEG_1643	Rv1165
MSMEG_0703	248	Rv0908	23	MSMEG_1644	Rv1170
MSMEG_0704	1	Rv0909	39	MSMEG_1645	Rv1176c
MSMEG_0709	111	Rv0920c	145	MSMEG_1647	Rv1185c
MSMEG_0711	18	Rv0924c	9	MSMEG_1648	Rv1191
MSMEG_0712	28	Rv0925c	78	MSMEG_1655	Rv1192
MSMEG_0716	447	Rv0934	312	MSMEG_1656	Rv1194c
MSMEG_0722	160	Rv0939	105	MSMEG_1660	Rv1201c
MSMEG_0726	20	Rv0941c	290	MSMEG_1663	Rv1202
MSMEG_0728	171	Rv0945	323	MSMEG_1665	Rv1205
MSMEG_0732	35	Rv0949	3	MSMEG_1666	Rv1207
MSMEG_0734	5	Rv0950c	75	MSMEG_1684	Rv1209
MSMEG_0735	1	Rv0951	68	MSMEG_1692	Rv1211
MSMEG_0736	1	Rv0954	49	MSMEG_1701	Rv1213
MSMEG_0739	369	Rv0957	162	MSMEG_1702	Rv1214c
MSMEG_0740	207	Rv0959	33	MSMEG_1708	Rv1215c
MSMEG_0752	96	Rv0965c	12	MSMEG_1716	Rv1219c
MSMEG_0753	28	Rv0966c	1	MSMEG_1720	Rv1220c

MSMEG_0758	133	Rv0971c	123	MSMEG_1725	Rv1221
MSMEG_0759	44	Rv0982	147	MSMEG_1734	Rv1241
MSMEG_0761	3	Rv0983	21	MSMEG_1735	Rv1244
MSMEG_0764	165	Rv0985c	46	MSMEG_1736	Rv1247c
MSMEG_0766	3	Rv0993	1	MSMEG_1739	Rv1259
MSMEG_0776	17	Rv0994	146	MSMEG_1740	Rv1261c
MSMEG_0781	90	Rv0996	73	MSMEG_1742	Rv1262c
MSMEG_0787	479	Rv0998	14	MSMEG_1745	Rv1263
MSMEG_0788	52	Rv1002c	58	MSMEG_1753	Rv1264
MSMEG_0791	98	Rv1006	123	MSMEG_1757	Rv1265
MSMEG_0796	423	Rv1007c	3	MSMEG_1809	Rv1276c
MSMEG_0797	16	Rv1008	436	MSMEG_1813	Rv1283c
MSMEG_0798	41	Rv1009	133	MSMEG_1817	Rv1289
MSMEG_0799	147	Rv1013	142	MSMEG_1818	Rv1302
MSMEG_0801	107	Rv1015c	129	MSMEG_1819	Rv1314c
MSMEG_0803	4	Rv1017c	32	MSMEG_1824	Rv1315
MSMEG_0806	402	Rv1018c	64	MSMEG_1826	Rv1322
MSMEG_0807	131	Rv1020	3	MSMEG_1830	Rv1322A
MSMEG_0814	110	Rv1021	59	MSMEG_1832	Rv1324
MSMEG_0818	30	Rv1023	59	MSMEG_1833	Rv1328
MSMEG_0826	165	Rv1028A	194	MSMEG_1845	Rv1330c
MSMEG_0828	76	Rv1033c	2	MSMEG_1852	Rv1331
MSMEG_0833	50	Rv1036c	1	MSMEG_1865	Rv1339
MSMEG_0839	71	Rv1042c	182	MSMEG_1867	Rv1340
MSMEG_0840	179	Rv1043c	105	MSMEG_1873	Rv1343c
MSMEG_0842	24	Rv1046c	52	MSMEG_1884	Rv1344
MSMEG_0844	312	Rv1054	363	MSMEG_1885	Rv1354c
MSMEG_0849	46	Rv1060	42	MSMEG_1890	Rv1360
MSMEG_0855	28	Rv1076	252	MSMEG_1893	Rv1377c
MSMEG_0859	39	Rv1081c	36	MSMEG_1900	Rv1398c
MSMEG_0861	1	Rv1085c	39	MSMEG_1901	Rv1399c
MSMEG_0862	90	Rv1089A	318	MSMEG_1913	Rv1400c
MSMEG_0866	19	Rv1116A	11	MSMEG_1920	Rv1401
MSMEG_0876	43	Rv1119c	224	MSMEG_1921	Rv1406
MSMEG_0880	139	Rv1128c	51	MSMEG_1922	Rv1408
MSMEG_0892	121	Rv1136	39	MSMEG_1926	Rv1411c
MSMEG_0894	57	Rv1140	71	MSMEG_1927	Rv1418
MSMEG_0895	53	Rv1159	99	MSMEG_1931	Rv1420
MSMEG_0896	393	Rv1166	165	MSMEG_1936	Rv1424c

MSMEG_0900	34	Rv1167c	270	MSMEG_1938	Rv1425
MSMEG_0901	215	Rv1168c	33	MSMEG_1940	Rv1440
MSMEG_0909	49	Rv1171	81	MSMEG_1942	Rv1441c
MSMEG_0911	66	Rv1178	186	MSMEG_1943	Rv1442
MSMEG_0912	65	Rv1179c	152	MSMEG_1944	Rv1443c
MSMEG_0913	96	Rv1181	183	MSMEG_1954	Rv1444c
MSMEG_0916	28	Rv1195	113	MSMEG_1957	Rv1454c
MSMEG_0918	35	Rv1198	23	MSMEG_1982	Rv1456c
MSMEG_0919	106	Rv1199c	374	MSMEG_1995	Rv1460
MSMEG_0932	287	Rv1200	343	MSMEG_2026	Rv1472
MSMEG_0940	246	Rv1203c	200	MSMEG_2027	Rv1486c
MSMEG_0941	389	Rv1204c	102	MSMEG_2032	Rv1497
MSMEG_0942	52	Rv1206	333	MSMEG_2037	Rv1516c
MSMEG_0944	78	Rv1208	398	MSMEG_2038	Rv1521
MSMEG_0945	39	Rv1210	98	MSMEG_2046	Rv1522c
MSMEG_0947	161	Rv1217c	342	MSMEG_2048	Rv1523
MSMEG_0948	43	Rv1223	39	MSMEG_2064	Rv1525
MSMEG_0951	329	Rv1227c	400	MSMEG_2083	Rv1530
MSMEG_0952	14	Rv1248c	88	MSMEG_2084	Rv1533
MSMEG_0962	37	Rv1251c	414	MSMEG_2086	Rv1538c
MSMEG_0965	135	Rv1252c	485	MSMEG_2089	Rv1539
MSMEG_0972	482	Rv1253	3	MSMEG_2127	Rv1543
MSMEG_0974	119	Rv1257c	1	MSMEG_2132	Rv1546
MSMEG_0975	27	Rv1258c	7	MSMEG_2136	Rv1549
MSMEG_0976	12	Rv1260	55	MSMEG_2137	Rv1556
MSMEG_0985	33	Rv1272c	92	MSMEG_2140	Rv1558
MSMEG_0986	33	Rv1280c	90	MSMEG_2144	Rv1559
MSMEG_0987	35	Rv1285	12	MSMEG_2153	Rv1560
MSMEG_0990	147	Rv1287	284	MSMEG_2176	Rv1564c
MSMEG_0991	327	Rv1300	66	MSMEG_2182	Rv1567c
MSMEG_0999	55	Rv1318c	3	MSMEG_2188	Rv1569
MSMEG_1002	107	Rv1319c	3	MSMEG_2192	Rv1593c
MSMEG_1003	259	Rv1320c	3	MSMEG_2200	Rv1608c
MSMEG_1005	24	Rv1321	35	MSMEG_2246	Rv1617
MSMEG_1007	11	Rv1323	71	MSMEG_2248	Rv1625c
MSMEG_1008	116	Rv1325c	445	MSMEG_2277	Rv1626
MSMEG_1012	40	Rv1326c	194	MSMEG_2290	Rv1629
MSMEG_1017	206	Rv1327c	46	MSMEG_2306	Rv1639c
MSMEG_1019	58	Rv1329c	135	MSMEG_2309	Rv1652

MSMEG_1023	88	Rv1336	3	MSMEG_2318	Rv1676
MSMEG_1025	65	Rv1337	26	MSMEG_2319	Rv1690
MSMEG_1027	68	Rv1338	195	MSMEG_2322	Rv1691
MSMEG_1028	26	Rv1345	94	MSMEG_2324	Rv1697
MSMEG_1030	63	Rv1346	95	MSMEG_2326	Rv1721c
MSMEG_1033	55	Rv1351	39	MSMEG_2329	Rv1722
MSMEG_1034	1	Rv1363c	162	MSMEG_2359	Rv1725c
MSMEG_1037	127	Rv1364c	68	MSMEG_2361	Rv1728c
MSMEG_1040	38	Rv1365c	71	MSMEG_2362	Rv1732c
MSMEG_1051	52	Rv1367c	74	MSMEG_2363	Rv1746
MSMEG_1052	46	Rv1368	208	MSMEG_2368	Rv1753c
MSMEG_1053	45	Rv1373	280	MSMEG_2370	Rv1820
MSMEG_1058	182	Rv1375	233	MSMEG_2371	Rv1834
MSMEG_1060	31	Rv1378c	85	MSMEG_2375	Rv1839c
MSMEG_1062	319	Rv1404	58	MSMEG_2377	Rv1847
MSMEG_1063	36	Rv1412	72	MSMEG_2378	Rv1857
MSMEG_1065	46	Rv1413	232	MSMEG_2381	Rv1862
MSMEG_1074	22	Rv1415	48	MSMEG_2382	Rv1872c
MSMEG_1076	42	Rv1416	205	MSMEG_2386	Rv1884c
MSMEG_1082	37	Rv1419	3	MSMEG_2392	Rv1890c
MSMEG_1089	103	Rv1421	15	MSMEG_2393	Rv1898
MSMEG_1112	46	Rv1423	131	MSMEG_2395	Rv1900c
MSMEG_1114	24	Rv1459c	94	MSMEG_2396	Rv1901
MSMEG_1116	36	Rv1463	90	MSMEG_2399	Rv1910c
MSMEG_1118	34	Rv1488	3	MSMEG_2408	Rv1911c
MSMEG_1134	49	Rv1491c	271	MSMEG_2410	Rv1913
MSMEG_1138	62	Rv1498A	42	MSMEG_2417	Rv1914c
MSMEG_1139	103	Rv1500	20	MSMEG_2421	Rv1920
MSMEG_1141	70	Rv1502	278	MSMEG_2423	Rv1925
MSMEG_1153	48	Rv1507A	498	MSMEG_2424	Rv1927
MSMEG_1165	53	Rv1508A	58	MSMEG_2428	Rv1932
MSMEG_1167	82	Rv1510	231	MSMEG_2430	Rv1942c
MSMEG_1168	51	Rv1515c	497	MSMEG_2432	Rv1944c
MSMEG_1169	65	Rv1518	105	MSMEG_2434	Rv1947
MSMEG_1174	44	Rv1519	27	MSMEG_2444	Rv1960c
MSMEG_1176	5	Rv1520	90	MSMEG_2445	Rv1962c
MSMEG_1177	66	Rv1535	468	MSMEG_2446	Rv1976c
MSMEG_1184	7	Rv1541c	3	MSMEG_2450	Rv1982c
MSMEG_1185	39	Rv1547	2	MSMEG_2468	Rv1984c

MSMEG_1188	45	Rv1548c	382	MSMEG_2516	Rv1987
MSMEG_1189	89	Rv1565c	4	MSMEG_2539	Rv1988
MSMEG_1191	32	Rv1568	334	MSMEG_2544	Rv2009
MSMEG_1193	307	Rv1577c	143	MSMEG_2546	Rv2024c
MSMEG_1195	26	Rv1599	18	MSMEG_2552	Rv2025c
MSMEG_1199	16	Rv1607	1	MSMEG_2556	Rv2027c
MSMEG_1200	44	Rv1610	196	MSMEG_2578	Rv2033c
MSMEG_1201	13	Rv1628c	341	MSMEG_2587	Rv2037c
MSMEG_1204	85	Rv1636	63	MSMEG_2589	Rv2045c
MSMEG_1205	56	Rv1638A	55	MSMEG_2590	Rv2046
MSMEG_1206	39	Rv1645c	8	MSMEG_2592	Rv2048c
MSMEG_1207	457	Rv1646	83	MSMEG_2601	Rv2054
MSMEG_1211	60	Rv1649	140	MSMEG_2605	Rv2061c
MSMEG_1212	59	Rv1658	69	MSMEG_2606	Rv2062c
MSMEG_1220	151	Rv1660	57	MSMEG_2612	Rv2064
MSMEG_1223	389	Rv1684	438	MSMEG_2618	Rv2065
MSMEG_1236	45	Rv1692	114	MSMEG_2621	Rv2068c
MSMEG_1238	367	Rv1700	227	MSMEG_2623	Rv2069
MSMEG_1240	182	Rv1703c	160	MSMEG_2641	Rv2073c
MSMEG_1242	247	Rv1706A	490	MSMEG_2647	Rv2077A
MSMEG_1245	65	Rv1706c	285	MSMEG_2652	Rv2080
MSMEG_1246	484	Rv1714	34	MSMEG_2653	Rv2088
MSMEG_1247	334	Rv1719	278	MSMEG_2658	Rv2097c
MSMEG_1251	40	Rv1730c	112	MSMEG_2660	Rv2104c
MSMEG_1253	317	Rv1731	33	MSMEG_2661	Rv2112c
MSMEG_1255	75	Rv1737c	201	MSMEG_2664	Rv2113
MSMEG_1257	107	Rv1740	31	MSMEG_2667	Rv2118c
MSMEG_1258	356	Rv1749c	150	MSMEG_2669	Rv2119
MSMEG_1260	21	Rv1754c	48	MSMEG_2670	Rv2122c
MSMEG_1262	301	Rv1765A	227	MSMEG_2676	Rv2124c
MSMEG_1263	243	Rv1765c	296	MSMEG_2683	Rv2125
MSMEG_1264	252	Rv1766	27	MSMEG_2684	Rv2129c
MSMEG_1267	449	Rv1770	72	MSMEG_2691	Rv2131c
MSMEG_1268	3	Rv1775	242	MSMEG_2692	Rv2142c
MSMEG_1270	71	Rv1779c	19	MSMEG_2698	Rv2149c
MSMEG_1272	147	Rv1780	2	MSMEG_2699	Rv2176
MSMEG_1274	41	Rv1785c	64	MSMEG_2700	Rv2178c
MSMEG_1275	37	Rv1793	17	MSMEG_2701	Rv2179c
MSMEG_1289	74	Rv1797	209	MSMEG_2728	Rv2192c

MSMEG_1292	39	Rv1801	18	MSMEG_2733	Rv2193
MSMEG_1295	83	Rv1803c	59	MSMEG_2735	Rv2201
MSMEG_1296	95	Rv1804c	56	MSMEG_2738	Rv2206
MSMEG_1300	47	Rv1810	240	MSMEG_2740	Rv2211c
MSMEG_1311	54	Rv1815	73	MSMEG_2744	Rv2213
MSMEG_1313	194	Rv1817	270	MSMEG_2748	Rv2217
MSMEG_1314	29	Rv1818c	75	MSMEG_2753	Rv2218
MSMEG_1316	3	Rv1819c	30	MSMEG_2754	Rv2219A
MSMEG_1317	3	Rv1826	46	MSMEG_2757	Rv2221c
MSMEG_1322	58	Rv1827	219	MSMEG_2761	Rv2222c
MSMEG_1339	64	Rv1829	102	MSMEG_2762	Rv2224c
MSMEG_1344	62	Rv1830	90	MSMEG_2764	Rv2227
MSMEG_1346	17	Rv1833c	17	MSMEG_2765	Rv2242
MSMEG_1350	37	Rv1837c	147	MSMEG_2772	Rv2243
MSMEG_1351	73	Rv1844c	57	MSMEG_2777	Rv2250A
MSMEG_1354	34	Rv1863c	1	MSMEG_2779	Rv2259
MSMEG_1356	159	Rv1878	492	MSMEG_2782	Rv2271
MSMEG_1361	319	Rv1895	138	MSMEG_2783	Rv2284
MSMEG_1364	240	Rv1902c	274	MSMEG_2784	Rv2285
MSMEG_1366	112	Rv1904	120	MSMEG_2791	Rv2289
MSMEG_1367	182	Rv1906c	1	MSMEG_2792	Rv2294
MSMEG_1369	21	Rv1922	2	MSMEG_2793	Rv2298
MSMEG_1370	50	Rv1923	12	MSMEG_2795	Rv2301
MSMEG_1375	34	Rv1924c	58	MSMEG_2819	Rv2305
MSMEG_1376	33	Rv1926c	55	MSMEG_2823	Rv2313c
MSMEG_1379	22	Rv1931c	275	MSMEG_2824	Rv2315c
MSMEG_1381	106	Rv1936	64	MSMEG_2839	Rv2326c
MSMEG_1392	32	Rv1941	47	MSMEG_2849	Rv2327
MSMEG_1394	37	Rv1948c	37	MSMEG_2868	Rv2343c
MSMEG_1397	112	Rv1951c	1	MSMEG_2897	Rv2345
MSMEG_1398	199	Rv1952	13	MSMEG_2902	Rv2362c
MSMEG_1401	79	Rv1955	471	MSMEG_2923	Rv2363
MSMEG_1402	30	Rv1956	175	MSMEG_2929	Rv2375
MSMEG_1403	43	Rv1986	48	MSMEG_2946	Rv2395
MSMEG_1404	161	Rv1991c	242	MSMEG_2957	Rv2400c
MSMEG_1405	107	Rv2000	24	MSMEG_2975	Rv2402
MSMEG_1407	33	Rv2002	34	MSMEG_2983	Rv2405
MSMEG_1410	20	Rv2012	51	MSMEG_3002	Rv2413c
MSMEG_1415	38	Rv2049c	247	MSMEG_3014	Rv2423

MSMEG_1419	1	Rv2053c	3	MSMEG_3016	Rv2445c
MSMEG_1423	301	Rv2078	19	MSMEG_3020	Rv2449c
MSMEG_1426	63	Rv2081c	134	MSMEG_3030	Rv2465c
MSMEG_1429	36	Rv2083	21	MSMEG_3032	Rv2467
MSMEG_1430	110	Rv2091c	31	MSMEG_3034	Rv2468c
MSMEG_1433	377	Rv2094c	20	MSMEG_3037	Rv2473
MSMEG_1434	180	Rv2114	256	MSMEG_3040	Rv2507
MSMEG_1435	277	Rv2116	152	MSMEG_3041	Rv2509
MSMEG_1440	301	Rv2126c	410	MSMEG_3059	Rv2510c
MSMEG_1442	159	Rv2127	104	MSMEG_3060	Rv2511
MSMEG_1446	304	Rv2128	170	MSMEG_3064	Rv2521
MSMEG_1448	92	Rv2132	16	MSMEG_3066	Rv2525c
MSMEG_1451	43	Rv2136c	372	MSMEG_3070	Rv2526
MSMEG_1455	186	Rv2138	85	MSMEG_3071	Rv2530A
MSMEG_1456	112	Rv2139	460	MSMEG_3076	Rv2536
MSMEG_1457	34	Rv2146c	62	MSMEG_3078	Rv2541
MSMEG_1465	135	Rv2153c	65	MSMEG_3098	Rv2547
MSMEG_1466	57	Rv2161c	31	MSMEG_3107	Rv2550c
MSMEG_1468	65	Rv2173	71	MSMEG_3117	Rv2556c
MSMEG_1470	351	Rv2174	361	MSMEG_3119	Rv2563
MSMEG_1471	97	Rv2181	363	MSMEG_3129	Rv2574
MSMEG_1472	15	Rv2183c	234	MSMEG_3144	Rv2576c
MSMEG_1473	58	Rv2184c	12	MSMEG_3157	Rv2581c
MSMEG_1479	80	Rv2190c	53	MSMEG_3162	Rv2582
MSMEG_1482	81	Rv2198c	91	MSMEG_3168	Rv2595
MSMEG_1483	126	Rv2200c	85	MSMEG_3171	Rv2601A
MSMEG_1487	293	Rv2202c	472	MSMEG_3173	Rv2616
MSMEG_1494	461	Rv2203	91	MSMEG_3174	Rv2636
MSMEG_1496	89	Rv2204c	96	MSMEG_3188	Rv2639c
MSMEG_1498	55	Rv2214c	39	MSMEG_3193	Rv2640c
MSMEG_1501	68	Rv2215	94	MSMEG_3197	Rv2671
MSMEG_1502	161	Rv2220	142	MSMEG_3199	Rv2672
MSMEG_1513	155	Rv2236c	340	MSMEG_3203	Rv2673
MSMEG_1514	105	Rv2239c	9	MSMEG_3216	Rv2674
MSMEG_1516	34	Rv2252	149	MSMEG_3217	Rv2676c
MSMEG_1518	95	Rv2255c	211	MSMEG_3223	Rv2679
MSMEG_1520	92	Rv2256c	79	MSMEG_3227	Rv2680
MSMEG_1521	5	Rv2270	315	MSMEG_3238	Rv2682c
MSMEG_1522	144	Rv2275	93	MSMEG_3239	Rv2697c

MSMEG_1523	400	Rv2280	40	MSMEG_3246	Rv2698
MSMEG_1524	292	Rv2283	375	MSMEG_3252	Rv2700
MSMEG_1525	257	Rv2291	221	MSMEG_3253	Rv2705c
MSMEG_1530	43	Rv2293c	79	MSMEG_3257	Rv2713
MSMEG_1531	95	Rv2295	5	MSMEG_3263	Rv2716
MSMEG_1540	108	Rv2299c	23	MSMEG_3264	Rv2717c
MSMEG_1542	79	Rv2300c	24	MSMEG_3278	Rv2720
MSMEG_1543	15	Rv2302	36	MSMEG_3284	Rv2738c
MSMEG_1544	282	Rv2304c	373	MSMEG_3286	Rv2740
MSMEG_1546	54	Rv2307D	118	MSMEG_3309	Rv2746c
MSMEG_1553	65	Rv2307c	480	MSMEG_3319	Rv2747
MSMEG_1556	121	Rv2308	256	MSMEG_3320	Rv2748c
MSMEG_1557	352	Rv2312	35	MSMEG_3345	Rv2758c
MSMEG_1558	416	Rv2314c	293	MSMEG_3358	Rv2778c
MSMEG_1559	430	Rv2320c	262	MSMEG_3360	Rv2788
MSMEG_1573	30	Rv2321c	243	MSMEG_3370	Rv2801c
MSMEG_1574	47	Rv2324	307	MSMEG_3376	Rv2809
MSMEG_1575	13	Rv2325c	130	MSMEG_3379	Rv2828c
MSMEG_1577	175	Rv2328	49	MSMEG_3400	Rv2830c
MSMEG_1580	335	Rv2329c	38	MSMEG_3408	Rv2831
MSMEG_1582	130	Rv2331	93	MSMEG_3411	Rv2847c
MSMEG_1583	148	Rv2338c	40	MSMEG_3413	Rv2861c
MSMEG_1585	201	Rv2339	182	MSMEG_3430	Rv2870c
MSMEG_1589	489	Rv2342	28	MSMEG_3433	Rv2871
MSMEG_1596	65	Rv2344c	1	MSMEG_3434	Rv2894c
MSMEG_1597	95	Rv2348c	66	MSMEG_3442	Rv2900c
MSMEG_1599	116	Rv2350c	57	MSMEG_3460	Rv2911
MSMEG_1602	40	Rv2351c	49	MSMEG_3464	Rv2914c
MSMEG_1603	122	Rv2352c	68	MSMEG_3468	Rv2916c
MSMEG_1604	102	Rv2356c	15	MSMEG_3475	Rv2922c
MSMEG_1608	422	Rv2357c	54	MSMEG_3479	Rv2923c
MSMEG_1609	28	Rv2358	213	MSMEG_3483	Rv2941
MSMEG_1618	132	Rv2360c	374	MSMEG_3488	Rv2951c
MSMEG_1620	53	Rv2361c	493	MSMEG_3491	Rv2952
MSMEG_1629	30	Rv2366c	264	MSMEG_3493	Rv2953
MSMEG_1632	104	Rv2367c	411	MSMEG_3499	Rv2954c
MSMEG_1633	488	Rv2368c	62	MSMEG_3500	Rv2956
MSMEG_1634	78	Rv2371	75	MSMEG_3506	Rv2960c
MSMEG_1636	1	Rv2373c	103	MSMEG_3513	Rv2981c

MSMEG_1637	224	Rv2374c	33	MSMEG_3514	Rv2982c
MSMEG_1638	134	Rv2376c	41	MSMEG_3515	Rv2989
MSMEG_1640	166	Rv2383c	36	MSMEG_3525	Rv2991
MSMEG_1642	1	Rv2385	264	MSMEG_3528	Rv2993c
MSMEG_1646	122	Rv2387	431	MSMEG_3532	Rv2996c
MSMEG_1652	320	Rv2388c	335	MSMEG_3535	Rv3003c
MSMEG_1654	38	Rv2390c	164	MSMEG_3558	Rv3004
MSMEG_1657	3	Rv2391	250	MSMEG_3561	Rv3005c
MSMEG_1659	73	Rv2394	161	MSMEG_3562	Rv3006
MSMEG_1661	19	Rv2401	65	MSMEG_3569	Rv3007c
MSMEG_1662	457	Rv2401A	18	MSMEG_3574	Rv3012c
MSMEG_1672	63	Rv2404c	231	MSMEG_3575	Rv3013
MSMEG_1673	116	Rv2406c	29	MSMEG_3592	Rv3014c
MSMEG_1675	42	Rv2412	29	MSMEG_3608	Rv3015c
MSMEG_1680	367	Rv2414c	89	MSMEG_3613	Rv3038c
MSMEG_1682	30	Rv2415c	186	MSMEG_3620	Rv3041c
MSMEG_1683	53	Rv2417c	182	MSMEG_3627	Rv3044
MSMEG_1685	226	Rv2421c	24	MSMEG_3628	Rv3046c
MSMEG_1689	487	Rv2428	40	MSMEG_3631	Rv3067
MSMEG_1698	82	Rv2431c	226	MSMEG_3634	Rv3068c
MSMEG_1717	5	Rv2436	35	MSMEG_3654	Rv3072c
MSMEG_1719	399	Rv2438c	3	MSMEG_3658	Rv3073c
MSMEG_1722	54	Rv2442c	134	MSMEG_3659	Rv3074
MSMEG_1723	110	Rv2443	35	MSMEG_3663	Rv3076
MSMEG_1726	158	Rv2444c	242	MSMEG_3670	Rv3091
MSMEG_1728	380	Rv2448c	30	MSMEG_3677	Rv3096
MSMEG_1729	8	Rv2457c	143	MSMEG_3681	Rv3099c
MSMEG_1732	68	Rv2461c	31	MSMEG_3682	Rv3105c
MSMEG_1733	281	Rv2477c	37	MSMEG_3708	Rv3137
MSMEG_1738	181	Rv2478c	68	MSMEG_3714	Rv3143
MSMEG_1741	1	Rv2484c	101	MSMEG_3748	Rv3170
MSMEG_1744	9	Rv2488c	198	MSMEG_3764	Rv3173c
MSMEG_1751	1	Rv2489c	6	MSMEG_3776	Rv3180c
MSMEG_1770	56	Rv2493	145	MSMEG_3778	Rv3190c
MSMEG_1773	179	Rv2495c	346	MSMEG_3780	Rv3194c
MSMEG_1778	268	Rv2496c	389	MSMEG_3799	Rv3198A
MSMEG_1782	212	Rv2497c	178	MSMEG_3804	Rv3198c
MSMEG_1786	134	Rv2499c	230	MSMEG_3807	Rv3199c
MSMEG_1791	42	Rv2500c	63	MSMEG_3810	Rv3200c

MSMEG_1799	53	Rv2501c	352	MSMEG_3812	Rv3202c
MSMEG_1801	471	Rv2502c	48	MSMEG_3819	Rv3210c
MSMEG_1802	54	Rv2504c	41	MSMEG_3839	Rv3214
MSMEG_1803	18	Rv2505c	54	MSMEG_3843	Rv3217c
MSMEG_1804	134	Rv2506	43	MSMEG_3845	Rv3218
MSMEG_1805	269	Rv2513	135	MSMEG_3846	Rv3224
MSMEG_1807	46	Rv2557	33	MSMEG_3848	Rv3258c
MSMEG_1812	171	Rv2558	6	MSMEG_3862	Rv3265c
MSMEG_1821	19	Rv2560	55	MSMEG_3863	Rv3271c
MSMEG_1822	125	Rv2561	3	MSMEG_3864	Rv3277
MSMEG_1828	125	Rv2565	180	MSMEG_3872	Rv3278c
MSMEG_1831	134	Rv2567	1	MSMEG_3879	Rv3279c
MSMEG_1834	26	Rv2569c	72	MSMEG_3881	Rv3280
MSMEG_1836	137	Rv2572c	21	MSMEG_3889	Rv3283
MSMEG_1838	32	Rv2573	159	MSMEG_3890	Rv3290c
MSMEG_1839	35	Rv2575	302	MSMEG_3906	Rv3291c
MSMEG_1843	162	Rv2579	26	MSMEG_3909	Rv3298c
MSMEG_1848	19	Rv2583c	13	MSMEG_3918	Rv3301c
MSMEG_1854	11	Rv2584c	432	MSMEG_3921	Rv3303c
MSMEG_1861	193	Rv2587c	37	MSMEG_3923	Rv3304
MSMEG_1862	158	Rv2588c	81	MSMEG_3961	Rv3306c
MSMEG_1863	228	Rv2589	35	MSMEG_3967	Rv3307
MSMEG_1868	32	Rv2590	59	MSMEG_3969	Rv3311
MSMEG_1869	76	Rv2594c	31	MSMEG_3978	Rv3312A
MSMEG_1870	32	Rv2597	56	MSMEG_3997	Rv3321c
MSMEG_1872	33	Rv2608	93	MSMEG_4007	Rv3329
MSMEG_1875	301	Rv2614c	38	MSMEG_4034	Rv3331
MSMEG_1880	55	Rv2621c	9	MSMEG_4129	Rv3337
MSMEG_1881	60	Rv2632c	59	MSMEG_4175	Rv3354
MSMEG_1882	34	Rv2633c	31	MSMEG_4183	Rv3369
MSMEG_1883	3	Rv2638	80	MSMEG_4186	Rv3372
MSMEG_1886	16	Rv2641	44	MSMEG_4188	Rv3390
MSMEG_1887	26	Rv2642	413	MSMEG_4189	Rv3392c
MSMEG_1895	27	Rv2651c	122	MSMEG_4190	Rv3393
MSMEG_1902	34	Rv2657c	6	MSMEG_4194	Rv3396c
MSMEG_1905	22	Rv2659c	31	MSMEG_4195	Rv3401
MSMEG_1914	90	Rv2661c	465	MSMEG_4200	Rv3407
MSMEG_1916	29	Rv2663	3	MSMEG_4214	Rv3409c
MSMEG_1917	26	Rv2665	134	MSMEG_4216	Rv3412

MSMEG_1918	282	Rv2678c	1	MSMEG_4239	Rv3421c
MSMEG_1919	89	Rv2696c	272	MSMEG_4242	Rv3436c
MSMEG_1929	403	Rv2707	274	MSMEG_4243	Rv3463
MSMEG_1930	223	Rv2712c	114	MSMEG_4244	Rv3464
MSMEG_1932	31	Rv2715	190	MSMEG_4258	Rv3489
MSMEG_1933	267	Rv2725c	183	MSMEG_4260	Rv3515c
MSMEG_1934	45	Rv2727c	33	MSMEG_4269	Rv3519
MSMEG_1935	58	Rv2729c	43	MSMEG_4273	Rv3547
MSMEG_1937	66	Rv2733c	9	MSMEG_4276	Rv3575c
MSMEG_1945	55	Rv2739c	1	MSMEG_4278	Rv3580c
MSMEG_1947	3	Rv2741	70	MSMEG_4285	Rv3592
MSMEG_1949	332	Rv2770c	52	MSMEG_4286	Rv3598c
MSMEG_1953	433	Rv2780	21	MSMEG_4287	Rv3604c
MSMEG_1955	115	Rv2793c	86	MSMEG_4289	Rv3623
MSMEG_1960	314	Rv2798c	27	MSMEG_4292	Rv3624c
MSMEG_1966	42	Rv2799	21	MSMEG_4295	Rv3630
MSMEG_1969	60	Rv2802c	173	MSMEG_4299	Rv3634c
MSMEG_1976	140	Rv2805	55	MSMEG_4304	Rv3645
MSMEG_1981	27	Rv2806	262	MSMEG_4308	Rv3647c
MSMEG_1985	62	Rv2808	264	MSMEG_4310	Rv3661
MSMEG_1986	41	Rv2818c	335	MSMEG_4322	Rv3667
MSMEG_1989	303	Rv2825c	40	MSMEG_4325	Rv3675
MSMEG_1991	20	Rv2837c	100	MSMEG_4330	Rv3678A
MSMEG_1996	8	Rv2839c	278	MSMEG_4333	Rv3682
MSMEG_1999	77	Rv2840c	58	MSMEG_4338	Rv3688c
MSMEG_2000	68	Rv2842c	91	MSMEG_4340	Rv3695
MSMEG_2002	107	Rv2851c	30	MSMEG_4353	Rv3698
MSMEG_2007	67	Rv2862c	50	MSMEG_4354	Rv3705A
MSMEG_2009	32	Rv2864c	1	MSMEG_4358	Rv3711c
MSMEG_2010	77	Rv2874	18	MSMEG_4381	Rv3712
MSMEG_2011	26	Rv2875	57	MSMEG_4384	Rv3719
MSMEG_2017	50	Rv2879c	209	MSMEG_4394	Rv3721c
MSMEG_2019	83	Rv2884	23	MSMEG_4447	Rv3732
MSMEG_2020	2	Rv2887	106	MSMEG_4454	Rv3744
MSMEG_2021	43	Rv2890c	195	MSMEG_4456	Rv3754
MSMEG_2029	21	Rv2892c	414	MSMEG_4464	Rv3755c
MSMEG_2033	22	Rv2898c	214	MSMEG_4474	Rv3759c
MSMEG_2039	19	Rv2905	134	MSMEG_4475	Rv3760
MSMEG_2042	37	Rv2912c	181	MSMEG_4485	Rv3761c

MSMEG_2043	3	Rv2917	297	MSMEG_4490	Rv3765c
MSMEG_2044	132	Rv2919c	98	MSMEG_4525	Rv3769
MSMEG_2045	169	Rv2927c	36	MSMEG_4534	Rv3772
MSMEG_2049	19	Rv2930	330	MSMEG_4544	Rv3773c
MSMEG_2072	123	Rv2936	293	MSMEG_4546	Rv3774
MSMEG_2074	172	Rv2937	332	MSMEG_4553	Rv3777
MSMEG_2075	55	Rv2942	160	MSMEG_4558	Rv3786c
MSMEG_2078	147	Rv2943	25	MSMEG_4563	Rv3793
MSMEG_2080	282	Rv2943A	30	MSMEG_4570	Rv3794
MSMEG_2081	105	Rv2945c	46	MSMEG_4581	Rv3811
MSMEG_2085	86	Rv2961	177	MSMEG_4601	Rv3816c
MSMEG_2087	91	Rv2964	13	MSMEG_4605	Rv3821
MSMEG_2091	26	Rv2979c	174	MSMEG_4622	Rv3826
MSMEG_2092	35	Rv2990c	155	MSMEG_4627	Rv3831
MSMEG_2103	148	Rv3025c	3	MSMEG_4630	Rv3832c
MSMEG_2105	131	Rv3029c	113	MSMEG_4632	Rv3833
MSMEG_2106	25	Rv3033	30	MSMEG_4633	Rv3834c
MSMEG_2107	46	Rv3045	43	MSMEG_4668	Rv3837c
MSMEG_2108	42	Rv3048c	40	MSMEG_4682	Rv3838c
MSMEG_2111	85	Rv3057c	304	MSMEG_4684	Rv3841
MSMEG_2116	48	Rv3060c	2	MSMEG_4687	Rv3843c
MSMEG_2117	20	Rv3075c	38	MSMEG_4690	Rv3854c
MSMEG_2124	79	Rv3080c	17	MSMEG_4691	Rv3888c
MSMEG_2125	45	Rv3093c	147	MSMEG_4693	Rv3905c
MSMEG_2128	3	Rv3095	3	MSMEG_4695	Rv3907c
MSMEG_2142	24	Rv3102c	3	MSMEG_4708	
MSMEG_2149	34	Rv3104c	282	MSMEG_4709	
MSMEG_2157	61	Rv3107c	214	MSMEG_4722	
MSMEG_2162	164	Rv3129	93	MSMEG_4739	
MSMEG_2164	78	Rv3132c	420	MSMEG_4756	
MSMEG_2173	17	Rv3138	9	MSMEG_4761	
MSMEG_2174	92	Rv3166c	62	MSMEG_4811	
MSMEG_2179	3	Rv3171c	115	MSMEG_4825	
MSMEG_2189	238	Rv3176c	68	MSMEG_4896	
MSMEG_2191	25	Rv3182	6	MSMEG_4901	
MSMEG_2195	62	Rv3186	241	MSMEG_4902	
MSMEG_2196	145	Rv3195	59	MSMEG_4910	
MSMEG_2197	361	Rv3196A	133	MSMEG_4911	
MSMEG_2198	103	Rv3197A	352	MSMEG_4915	

MSMEG_2199	72	Rv3203	112	MSMEG_4921	
MSMEG_2201	59	Rv3207c	171	MSMEG_4923	
MSMEG_2202	18	Rv3208A	38	MSMEG_4932	
MSMEG_2204	131	Rv3209	177	MSMEG_4934	
MSMEG_2205	38	Rv3211	189	MSMEG_4959	
MSMEG_2206	47	Rv3221c	65	MSMEG_4967	
MSMEG_2211	17	Rv3233c	455	MSMEG_4974	
MSMEG_2224	19	Rv3239c	183	MSMEG_4976	
MSMEG_2225	124	Rv3240c	39	MSMEG_4980	
MSMEG_2226	427	Rv3243c	15	MSMEG_4985	
MSMEG_2227	41	Rv3259	84	MSMEG_4990	
MSMEG_2245	17	Rv3260c	235	MSMEG_4999	
MSMEG_2259	36	Rv3261	87	MSMEG_5007	
MSMEG_2261	29	Rv3266c	59	MSMEG_5008	
MSMEG_2268	149	Rv3268	6	MSMEG_5010	
MSMEG_2272	29	Rv3269	46	MSMEG_5011	
MSMEG_2273	54	Rv3272	32	MSMEG_5021	
MSMEG_2280	53	Rv3274c	30	MSMEG_5022	
MSMEG_2282	107	Rv3275c	106	MSMEG_5025	
MSMEG_2284	40	Rv3276c	219	MSMEG_5028	
MSMEG_2287	11	Rv3284	349	MSMEG_5029	
MSMEG_2288	116	Rv3285	57	MSMEG_5030	
MSMEG_2289	54	Rv3287c	18	MSMEG_5037	
MSMEG_2292	40	Rv3288c	3	MSMEG_5038	
MSMEG_2297	206	Rv3292	289	MSMEG_5043	
MSMEG_2303	88	Rv3309c	138	MSMEG_5044	
MSMEG_2307	68	Rv3313c	260	MSMEG_5047	
MSMEG_2308	26	Rv3322c	302	MSMEG_5052	
MSMEG_2310	63	Rv3323c	12	MSMEG_5055	
MSMEG_2313	55	Rv3324c	95	MSMEG_5064	
MSMEG_2314	1	Rv3330	1	MSMEG_5078	
MSMEG_2317	127	Rv3339c	7	MSMEG_5079	
MSMEG_2320	38	Rv3360	1	MSMEG_5083	
MSMEG_2331	52	Rv3364c	145	MSMEG_5085	
MSMEG_2332	46	Rv3366	481	MSMEG_5088	
MSMEG_2333	45	Rv3371	30	MSMEG_5096	
MSMEG_2338	182	Rv3379c	252	MSMEG_5103	
MSMEG_2349	435	Rv3382c	9	MSMEG_5104	
MSMEG_2351	102	Rv3386	33	MSMEG_5114	

MSMEG_2352	72	Rv3388	18	MSMEG_5116	
MSMEG_2357	3	Rv3400	19	MSMEG_5120	
MSMEG_2366	26	Rv3402c	116	MSMEG_5132	
MSMEG_2372	210	Rv3403c	30	MSMEG_5150	
MSMEG_2373	30	Rv3404c	50	MSMEG_5156	
MSMEG_2387	111	Rv3411c	34	MSMEG_5164	
MSMEG_2389	39	Rv3416	95	MSMEG_5169	
MSMEG_2397	41	Rv3418c	170	MSMEG_5172	
MSMEG_2398	69	Rv3419c	275	MSMEG_5174	
MSMEG_2400	63	Rv3426	328	MSMEG_5175	
MSMEG_2403	9	Rv3430c	472	MSMEG_5180	
MSMEG_2405	51	Rv3435c	30	MSMEG_5181	
MSMEG_2407	152	Rv3437	78	MSMEG_5185	
MSMEG_2409	121	Rv3440c	126	MSMEG_5187	
MSMEG_2412	45	Rv3443c	119	MSMEG_5190	
MSMEG_2416	26	Rv3451	42	MSMEG_5191	
MSMEG_2420	23	Rv3455c	310	MSMEG_5193	
MSMEG_2425	43	Rv3462c	155	MSMEG_5194	
MSMEG_2426	16	Rv3466	78	MSMEG_5203	
MSMEG_2427	292	Rv3482c	126	MSMEG_5210	
MSMEG_2435	113	Rv3486	75	MSMEG_5213	
MSMEG_2439	78	Rv3488	30	MSMEG_5215	
MSMEG_2440	122	Rv3493c	65	MSMEG_5217	
MSMEG_2442	87	Rv3510c	3	MSMEG_5222	
MSMEG_2443	82	Rv3514	59	MSMEG_5223	
MSMEG_2449	29	Rv3516	176	MSMEG_5225	
MSMEG_2451	446	Rv3517	403	MSMEG_5238	
MSMEG_2453	20	Rv3522	352	MSMEG_5239	
MSMEG_2455	58	Rv3524	12	MSMEG_5244	
MSMEG_2456	27	Rv3526	28	MSMEG_5248	
MSMEG_2467	74	Rv3528c	58	MSMEG_5249	
MSMEG_2469	32	Rv3533c	28	MSMEG_5252	
MSMEG_2472	33	Rv3536c	156	MSMEG_5256	
MSMEG_2475	28	Rv3543c	431	MSMEG_5258	
MSMEG_2478	58	Rv3544c	258	MSMEG_5267	
MSMEG_2486	97	Rv3545c	21	MSMEG_5270	
MSMEG_2487	44	Rv3546	199	MSMEG_5272	
MSMEG_2488	41	Rv3550	36	MSMEG_5276	
MSMEG_2490	151	Rv3554	69	MSMEG_5281	

MSMEG_2495	445	Rv3557c	270	MSMEG_5285	
MSMEG_2497	44	Rv3558	61	MSMEG_5291	
MSMEG_2506	217	Rv3568c	64	MSMEG_5332	
MSMEG_2507	66	Rv3570c	29	MSMEG_5340	
MSMEG_2510	412	Rv3572	33	MSMEG_5351	
MSMEG_2511	63	Rv3574	196	MSMEG_5362	
MSMEG_2512	302	Rv3576	24	MSMEG_5395	
MSMEG_2517	51	Rv3577	66	MSMEG_5404	
MSMEG_2519	141	Rv3590c	30	MSMEG_5405	
MSMEG_2520	96	Rv3596c	55	MSMEG_5424	
MSMEG_2526	80	Rv3597c	46	MSMEG_5440	
MSMEG_2527	25	Rv3600c	209	MSMEG_5445	
MSMEG_2528	61	Rv3603c	61	MSMEG_5446	
MSMEG_2529	24	Rv3607c	263	MSMEG_5447	
MSMEG_2531	100	Rv3610c	48	MSMEG_5448	
MSMEG_2532	159	Rv3611	262	MSMEG_5451	
MSMEG_2540	57	Rv3616c	65	MSMEG_5454	
MSMEG_2541	62	Rv3631	2	MSMEG_5458	
MSMEG_2542	106	Rv3633	111	MSMEG_5471	
MSMEG_2543	71	Rv3635	58	MSMEG_5513	
MSMEG_2545	24	Rv3639c	96	MSMEG_5516	
MSMEG_2553	312	Rv3642c	19	MSMEG_5519	
MSMEG_2570	206	Rv3644c	299	MSMEG_5520	
MSMEG_2571	38	Rv3660c	326	MSMEG_5527	
MSMEG_2572	449	Rv3673c	320	MSMEG_5538	
MSMEG_2574	123	Rv3681c	144	MSMEG_5541	
MSMEG_2579	83	Rv3696c	1	MSMEG_5544	
MSMEG_2581	93	Rv3697c	221	MSMEG_5546	
MSMEG_2582	50	Rv3724A	25	MSMEG_5548	
MSMEG_2584	1	Rv3730c	5	MSMEG_5562	
MSMEG_2585	3	Rv3737	1	MSMEG_5566	
MSMEG_2594	1	Rv3753c	21	MSMEG_5567	
MSMEG_2595	92	Rv3766	166	MSMEG_5570	
MSMEG_2597	298	Rv3767c	35	MSMEG_5577	
MSMEG_2604	353	Rv3768	29	MSMEG_5608	
MSMEG_2611	41	Rv3778c	71	MSMEG_5615	
MSMEG_2613	74	Rv3780	91	MSMEG_5635	
MSMEG_2624	59	Rv3785	43	MSMEG_5637	
MSMEG_2625	120	Rv3788	127	MSMEG_5642	

MSMEG_2630	47	Rv3789	64	MSMEG_5645	
MSMEG_2635	65	Rv3790	1	MSMEG_5648	
MSMEG_2639	15	Rv3801c	177	MSMEG_5650	
MSMEG_2643	263	Rv3819	66	MSMEG_5654	
MSMEG_2645	66	Rv3825c	10	MSMEG_5662	
MSMEG_2654	73	Rv3828c	15	MSMEG_5664	
MSMEG_2655	327	Rv3829c	10	MSMEG_5666	
MSMEG_2656	258	Rv3830c	499	MSMEG_5669	
MSMEG_2657	215	Rv3835	1	MSMEG_5670	
MSMEG_2659	27	Rv3840	395	MSMEG_5673	
MSMEG_2663	477	Rv3846	85	MSMEG_5675	
MSMEG_2668	182	Rv3847	44	MSMEG_5679	
MSMEG_2672	27	Rv3848	151	MSMEG_5688	
MSMEG_2673	399	Rv3849	51	MSMEG_5689	
MSMEG_2675	5	Rv3850	45	MSMEG_5693	
MSMEG_2679	382	Rv3852	50	MSMEG_5712	
MSMEG_2688	40	Rv3857c	3	MSMEG_5715	
MSMEG_2693	62	Rv3859c	187	MSMEG_5717	
MSMEG_2694	36	Rv3864	61	MSMEG_5718	
MSMEG_2695	70	Rv3865	19	MSMEG_5724	
MSMEG_2696	47	Rv3883c	77	MSMEG_5734	
MSMEG_2720	76	Rv3899c	293	MSMEG_5736	
MSMEG_2722	23	Rv3901c	481	MSMEG_5738	
MSMEG_2723	91	Rv3902c	144	MSMEG_5739	
MSMEG_2727	30	Rv3906c	14	MSMEG_5752	
MSMEG_2731	37	Rv3908	259	MSMEG_5753	
MSMEG_2732	34	Rv3909	213	MSMEG_5766	
MSMEG_2736	4	Rv3912	432	MSMEG_5776	
MSMEG_2737	29			MSMEG_5777	
MSMEG_2739	44			MSMEG_5791	
MSMEG_2742	85			MSMEG_5795	
MSMEG_2743	40			MSMEG_5796	
MSMEG_2745	215			MSMEG_5817	
MSMEG_2746	59			MSMEG_5819	
MSMEG_2750	23			MSMEG_5824	
MSMEG_2751	197			MSMEG_5827	
MSMEG_2752	28			MSMEG_5829	
MSMEG_2755	3			MSMEG_5837	
MSMEG_2756	50			MSMEG_5838	

MSMEG_2758	123			MSMEG_5841	
MSMEG_2760	77			MSMEG_5850	
MSMEG_2763	26			MSMEG_5852	
MSMEG_2768	46			MSMEG_5860	
MSMEG_2770	69			MSMEG_5873	
MSMEG_2771	48			MSMEG_5877	
MSMEG_2773	87			MSMEG_5888	
MSMEG_2774	25			MSMEG_5890	
MSMEG_2775	56			MSMEG_5891	
MSMEG_2778	98			MSMEG_5901	
MSMEG_2780	1			MSMEG_5914	
MSMEG_2785	402			MSMEG_5941	
MSMEG_2794	28			MSMEG_5996	
MSMEG_2798	33			MSMEG_6000	
MSMEG_2799	181			MSMEG_6009	
MSMEG_2800	52			MSMEG_6012	
MSMEG_2804	31			MSMEG_6043	
MSMEG_2813	2			MSMEG_6053	
MSMEG_2816	29			MSMEG_6054	
MSMEG_2820	392			MSMEG_6059	
MSMEG_2822	432			MSMEG_6060	
MSMEG_2826	28			MSMEG_6074	
MSMEG_2829	107			MSMEG_6079	
MSMEG_2840	3			MSMEG_6081	
MSMEG_2843	145			MSMEG_6082	
MSMEG_2881	17			MSMEG_6084	
MSMEG_2890	328			MSMEG_6086	
MSMEG_2895	34			MSMEG_6094	
MSMEG_2898	36			MSMEG_6106	
MSMEG_2899	24			MSMEG_6125	
MSMEG_2900	29			MSMEG_6135	
MSMEG_2901	49			MSMEG_6137	
MSMEG_2903	192			MSMEG_6143	
MSMEG_2911	66			MSMEG_6144	
MSMEG_2919	62			MSMEG_6150	
MSMEG_2931	35			MSMEG_6154	
MSMEG_2932	185			MSMEG_6174	
MSMEG_2934	84			MSMEG_6187	
MSMEG_2936	173			MSMEG_6188	

MSMEG_2937	23			MSMEG_6192	
MSMEG_2940	32			MSMEG_6193	
MSMEG_2941	94			MSMEG_6202	
MSMEG_2942	35			MSMEG_6208	
MSMEG_2943	24			MSMEG_6209	
MSMEG_2947	53			MSMEG_6211	
MSMEG_2953	19			MSMEG_6214	
MSMEG_2956	183			MSMEG_6226	
MSMEG_2959	63			MSMEG_6228	
MSMEG_2960	67			MSMEG_6234	
MSMEG_2961	34			MSMEG_6253	
MSMEG_2964	74			MSMEG_6258	
MSMEG_2965	26			MSMEG_6275	
MSMEG_2973	175			MSMEG_6276	
MSMEG_2974	44			MSMEG_6283	
MSMEG_2976	174			MSMEG_6285	
MSMEG_2982	86			MSMEG_6290	
MSMEG_2989	73			MSMEG_6293	
MSMEG_2994	35			MSMEG_6294	
MSMEG_3000	42			MSMEG_6306	
MSMEG_3003	29			MSMEG_6307	
MSMEG_3008	81			MSMEG_6313	
MSMEG_3013	57			MSMEG_6315	
MSMEG_3015	58			MSMEG_6317	
MSMEG_3017	1			MSMEG_6328	
MSMEG_3019	40			MSMEG_6329	
MSMEG_3025	118			MSMEG_6330	
MSMEG_3029	448			MSMEG_6336	
MSMEG_3033	17			MSMEG_6337	
MSMEG_3035	211			MSMEG_6338	
MSMEG_3042	302			MSMEG_6343	
MSMEG_3045	97			MSMEG_6345	
MSMEG_3046	489			MSMEG_6351	
MSMEG_3050	74			MSMEG_6361	
MSMEG_3053	125			MSMEG_6362	
MSMEG_3055	174			MSMEG_6368	
MSMEG_3056	25			MSMEG_6370	
MSMEG_3057	443			MSMEG_6372	
MSMEG_3062	55			MSMEG_6387	

MSMEG_3065	216			MSMEG_6388	
MSMEG_3068	54			MSMEG_6395	
MSMEG_3072	77			MSMEG_6406	
MSMEG_3079	157			MSMEG_6409	
MSMEG_3080	171			MSMEG_6413	
MSMEG_3081	290			MSMEG_6418	
MSMEG_3082	40			MSMEG_6420	
MSMEG_3084	33			MSMEG_6422	
MSMEG_3085	156			MSMEG_6439	
MSMEG_3086	260			MSMEG_6440	
MSMEG_3087	42			MSMEG_6441	
MSMEG_3095	92			MSMEG_6454	
MSMEG_3103	48			MSMEG_6464	
MSMEG_3111	377			MSMEG_6465	
MSMEG_3121	37			MSMEG_6468	
MSMEG_3122	28			MSMEG_6475	
MSMEG_3132	314			MSMEG_6479	
MSMEG_3138	33			MSMEG_6485	
MSMEG_3140	148			MSMEG_6499	
MSMEG_3141	30			MSMEG_6502	
MSMEG_3143	45			MSMEG_6503	
MSMEG_3145	97			MSMEG_6507	
MSMEG_3147	246			MSMEG_6520	
MSMEG_3150	27			MSMEG_6526	
MSMEG_3153	3			MSMEG_6546	
MSMEG_3155	290			MSMEG_6547	
MSMEG_3159	121			MSMEG_6557	
MSMEG_3160	499			MSMEG_6559	
MSMEG_3161	349			MSMEG_6571	
MSMEG_3163	398			MSMEG_6574	
MSMEG_3165	107			MSMEG_6580	
MSMEG_3169	231			MSMEG_6583	
MSMEG_3178	2			MSMEG_6584	
MSMEG_3180	27			MSMEG_6598	
MSMEG_3183	47			MSMEG_6605	
MSMEG_3186	79			MSMEG_6617	
MSMEG_3187	7			MSMEG_6620	
MSMEG_3189	123			MSMEG_6627	
MSMEG_3191	37			MSMEG_6639	

MSMEG_3194	42			MSMEG_6643	
MSMEG_3196	254			MSMEG_6661	
MSMEG_3204	126			MSMEG_6666	
MSMEG_3215	79			MSMEG_6685	
MSMEG_3219	49			MSMEG_6729	
MSMEG_3225	56			MSMEG_6731	
MSMEG_3228	57			MSMEG_6733	
MSMEG_3233	32			MSMEG_6738	
MSMEG_3235	122			MSMEG_6741	
MSMEG_3236	278			MSMEG_6745	
MSMEG_3237	99			MSMEG_6746	
MSMEG_3247	121			MSMEG_6749	
MSMEG_3248	433			MSMEG_6752	
MSMEG_3249	20			MSMEG_6753	
MSMEG_3250	32			MSMEG_6757	
MSMEG_3256	377			MSMEG_6776	
MSMEG_3259	83			MSMEG_6778	
MSMEG_3265	25			MSMEG_6779	
MSMEG_3275	210			MSMEG_6780	
MSMEG_3280	410			MSMEG_6807	
MSMEG_3281	16			MSMEG_6815	
MSMEG_3283	43			MSMEG_6828	
MSMEG_3287	45			MSMEG_6829	
MSMEG_3292	243			MSMEG_6839	
MSMEG_3297	83			MSMEG_6857	
MSMEG_3304	342			MSMEG_6862	
MSMEG_3306	136			MSMEG_6869	
MSMEG_3308	27			MSMEG_6882	
MSMEG_3312	198			MSMEG_6886	
MSMEG_3314	15			MSMEG_6902	
MSMEG_3317	34			MSMEG_6906	
MSMEG_3323	88			MSMEG_6907	
MSMEG_3325	131			MSMEG_6909	
MSMEG_3327	266			MSMEG_6913	
MSMEG_3332	132			MSMEG_6916	
MSMEG_3335	3			MSMEG_6917	
MSMEG_3336	198			MSMEG_6921	
MSMEG_3337	216			MSMEG_6926	
MSMEG_3338	480			MSMEG_6927	

MSMEG_3341	199				
MSMEG_3342	75				
MSMEG_3344	1				
MSMEG_3346	50				
MSMEG_3352	126				
MSMEG_3357	25				
MSMEG_3359	136				
MSMEG_3361	70				
MSMEG_3363	27				
MSMEG_3365	220				
MSMEG_3375	43				
MSMEG_3377	297				
MSMEG_3378	50				
MSMEG_3380	27				
MSMEG_3382	12				
MSMEG_3383	335				
MSMEG_3389	24				
MSMEG_3395	181				
MSMEG_3407	16				
MSMEG_3412	30				
MSMEG_3418	254				
MSMEG_3419	41				
MSMEG_3421	85				
MSMEG_3422	96				
MSMEG_3425	175				
MSMEG_3426	34				
MSMEG_3432	279				
MSMEG_3435	23				
MSMEG_3437	49				
MSMEG_3438	413				
MSMEG_3439	49				
MSMEG_3440	96				
MSMEG_3441	98				
MSMEG_3443	34				
MSMEG_3447	53				
MSMEG_3449	86				
MSMEG_3451	34				
MSMEG_3452	67				
MSMEG_3454	4				

MSMEG_3457	19				
MSMEG_3459	61				
MSMEG_3466	273				
MSMEG_3471	37				
MSMEG_3472	72				
MSMEG_3476	26				
MSMEG_3477	68				
MSMEG_3480	1				
MSMEG_3487	153				
MSMEG_3489	123				
MSMEG_3490	98				
MSMEG_3494	28				
MSMEG_3495	124				
MSMEG_3497	34				
MSMEG_3501	63				
MSMEG_3503	134				
MSMEG_3504	34				
MSMEG_3505	65				
MSMEG_3507	75				
MSMEG_3511	62				
MSMEG_3519	55				
MSMEG_3520	30				
MSMEG_3531	95				
MSMEG_3534	38				
MSMEG_3538	102				
MSMEG_3540	30				
MSMEG_3545	119				
MSMEG_3564	51				
MSMEG_3566	35				
MSMEG_3570	341				
MSMEG_3578	70				
MSMEG_3580	73				
MSMEG_3583	77				
MSMEG_3589	14				
MSMEG_3594	2				
MSMEG_3605	100				
MSMEG_3612	21				
MSMEG_3614	45				
MSMEG_3615	297				

MSMEG_3616	71				
MSMEG_3619	49				
MSMEG_3621	36				
MSMEG_3630	32				
MSMEG_3632	20				
MSMEG_3633	343				
MSMEG_3636	38				
MSMEG_3639	430				
MSMEG_3641	35				
MSMEG_3642	286				
MSMEG_3644	152				
MSMEG_3645	76				
MSMEG_3647	200				
MSMEG_3648	31				
MSMEG_3649	87				
MSMEG_3650	227				
MSMEG_3660	46				
MSMEG_3666	1				
MSMEG_3667	50				
MSMEG_3668	127				
MSMEG_3671	1				
MSMEG_3674	22				
MSMEG_3675	27				
MSMEG_3678	48				
MSMEG_3680	102				
MSMEG_3683	3				
MSMEG_3685	27				
MSMEG_3686	48				
MSMEG_3687	22				
MSMEG_3689	29				
MSMEG_3690	53				
MSMEG_3692	44				
MSMEG_3693	51				
MSMEG_3694	38				
MSMEG_3696	107				
MSMEG_3706	21				
MSMEG_3718	26				
MSMEG_3725	166				
MSMEG_3727	172				

MSMEG_3730	96				
MSMEG_3733	38				
MSMEG_3736	47				
MSMEG_3737	47				
MSMEG_3741	96				
MSMEG_3743	81				
MSMEG_3744	482				
MSMEG_3746	61				
MSMEG_3749	252				
MSMEG_3750	134				
MSMEG_3752	3				
MSMEG_3754	141				
MSMEG_3759	1				
MSMEG_3765	109				
MSMEG_3766	154				
MSMEG_3767	42				
MSMEG_3768	3				
MSMEG_3774	271				
MSMEG_3775	261				
MSMEG_3784	35				
MSMEG_3785	45				
MSMEG_3786	67				
MSMEG_3791	64				
MSMEG_3792	31				
MSMEG_3794	192				
MSMEG_3795	37				
MSMEG_3796	374				
MSMEG_3800	73				
MSMEG_3802	146				
MSMEG_3803	43				
MSMEG_3808	73				
MSMEG_3811	30				
MSMEG_3831	456				
MSMEG_3833	124				
MSMEG_3835	239				
MSMEG_3836	27				
MSMEG_3854	140				
MSMEG_3855	273				
MSMEG_3856	58				

MSMEG_3858	64				
MSMEG_3861	101				
MSMEG_3869	410				
MSMEG_3874	23				
MSMEG_3877	158				
MSMEG_3880	1				
MSMEG_3887	21				
MSMEG_3896	49				
MSMEG_3901	58				
MSMEG_3902	38				
MSMEG_3903	25				
MSMEG_3904	184				
MSMEG_3916	128				
MSMEG_3919	240				
MSMEG_3927	2				
MSMEG_3930	127				
MSMEG_3950	27				
MSMEG_3958	68				
MSMEG_3959	50				
MSMEG_3960	120				
MSMEG_3962	110				
MSMEG_3965	220				
MSMEG_3968	32				
MSMEG_3974	56				
MSMEG_3975	38				
MSMEG_3976	24				
MSMEG_3977	51				
MSMEG_3979	22				
MSMEG_3980	57				
MSMEG_3988	36				
MSMEG_4022	20				
MSMEG_4024	5				
MSMEG_4029	44				
MSMEG_4030	63				
MSMEG_4037	289				
MSMEG_4038	102				
MSMEG_4041	59				
MSMEG_4042	106				
MSMEG_4044	26				

MSMEG_4052	254				
MSMEG_4053	69				
MSMEG_4055	185				
MSMEG_4057	69				
MSMEG_4067	54				
MSMEG_4068	326				
MSMEG_4071	105				
MSMEG_4073	19				
MSMEG_4079	122				
MSMEG_4081	286				
MSMEG_4085	165				
MSMEG_4087	143				
MSMEG_4088	103				
MSMEG_4090	77				
MSMEG_4108	338				
MSMEG_4109	302				
MSMEG_4113	214				
MSMEG_4116	298				
MSMEG_4120	170				
MSMEG_4121	22				
MSMEG_4123	78				
MSMEG_4125	20				
MSMEG_4126	39				
MSMEG_4128	41				
MSMEG_4130	36				
MSMEG_4132	145				
MSMEG_4133	158				
MSMEG_4134	291				
MSMEG_4142	1				
MSMEG_4143	54				
MSMEG_4145	51				
MSMEG_4151	207				
MSMEG_4153	64				
MSMEG_4155	299				
MSMEG_4159	15				
MSMEG_4160	308				
MSMEG_4161	251				
MSMEG_4163	342				
MSMEG_4165	408				

MSMEG_4166	469				
MSMEG_4174	31				
MSMEG_4179	23				
MSMEG_4180	266				
MSMEG_4185	91				
MSMEG_4192	3				
MSMEG_4196	49				
MSMEG_4197	143				
MSMEG_4199	47				
MSMEG_4215	36				
MSMEG_4218	96				
MSMEG_4219	68				
MSMEG_4222	100				
MSMEG_4225	90				
MSMEG_4226	117				
MSMEG_4228	244				
MSMEG_4230	244				
MSMEG_4232	495				
MSMEG_4235	312				
MSMEG_4236	76				
MSMEG_4237	112				
MSMEG_4238	344				
MSMEG_4240	58				
MSMEG_4245	372				
MSMEG_4246	257				
MSMEG_4248	35				
MSMEG_4249	89				
MSMEG_4251	54				
MSMEG_4252	1				
MSMEG_4253	30				
MSMEG_4256	339				
MSMEG_4257	37				
MSMEG_4261	422				
MSMEG_4264	167				
MSMEG_4268	51				
MSMEG_4270	265				
MSMEG_4271	27				
MSMEG_4272	80				
MSMEG_4274	239				

MSMEG_4280	30				
MSMEG_4282	29				
MSMEG_4283	93				
MSMEG_4284	149				
MSMEG_4290	26				
MSMEG_4298	69				
MSMEG_4302	38				
MSMEG_4307	200				
MSMEG_4313	36				
MSMEG_4314	12				
MSMEG_4320	390				
MSMEG_4321	74				
MSMEG_4323	59				
MSMEG_4326	333				
MSMEG_4336	90				
MSMEG_4337	25				
MSMEG_4347	47				
MSMEG_4349	14				
MSMEG_4351	34				
MSMEG_4359	481				
MSMEG_4360	59				
MSMEG_4361	46				
MSMEG_4365	55				
MSMEG_4366	73				
MSMEG_4367	244				
MSMEG_4368	33				
MSMEG_4371	313				
MSMEG_4379	10				
MSMEG_4383	60				
MSMEG_4385	81				
MSMEG_4392	104				
MSMEG_4395	45				
MSMEG_4396	41				
MSMEG_4399	218				
MSMEG_4401	11				
MSMEG_4402	107				
MSMEG_4410	72				
MSMEG_4430	142				
MSMEG_4431	14				

MSMEG_4449	49				
MSMEG_4450	35				
MSMEG_4453	330				
MSMEG_4458	31				
MSMEG_4460	455				
MSMEG_4461	54				
MSMEG_4476	65				
MSMEG_4479	26				
MSMEG_4480	173				
MSMEG_4481	239				
MSMEG_4483	1				
MSMEG_4486	19				
MSMEG_4497	57				
MSMEG_4499	48				
MSMEG_4504	166				
MSMEG_4507	53				
MSMEG_4518	30				
MSMEG_4527	220				
MSMEG_4533	207				
MSMEG_4536	211				
MSMEG_4537	34				
MSMEG_4550	192				
MSMEG_4555	407				
MSMEG_4556	9				
MSMEG_4559	326				
MSMEG_4561	139				
MSMEG_4564	84				
MSMEG_4565	364				
MSMEG_4571	70				
MSMEG_4580	78				
MSMEG_4595	170				
MSMEG_4598	374				
MSMEG_4599	259				
MSMEG_4604	78				
MSMEG_4609	12				
MSMEG_4614	341				
MSMEG_4617	1				
MSMEG_4618	38				
MSMEG_4620	1				

MSMEG_4625	110				
MSMEG_4626	236				
MSMEG_4628	355				
MSMEG_4631	39				
MSMEG_4634	27				
MSMEG_4640	96				
MSMEG_4643	225				
MSMEG_4645	144				
MSMEG_4646	118				
MSMEG_4647	90				
MSMEG_4654	30				
MSMEG_4655	58				
MSMEG_4657	53				
MSMEG_4658	67				
MSMEG_4659	46				
MSMEG_4661	61				
MSMEG_4671	138				
MSMEG_4673	44				
MSMEG_4678	27				
MSMEG_4679	475				
MSMEG_4688	33				
MSMEG_4689	47				
MSMEG_4694	75				
MSMEG_4698	109				
MSMEG_4699	104				
MSMEG_4700	77				
MSMEG_4705	37				
MSMEG_4712	155				
MSMEG_4717	31				
MSMEG_4718	41				
MSMEG_4720	5				
MSMEG_4723	219				
MSMEG_4727	25				
MSMEG_4735	1				
MSMEG_4737	289				
MSMEG_4738	154				
MSMEG_4740	53				
MSMEG_4741	31				
MSMEG_4742	59				

MSMEG_4743	58				
MSMEG_4748	79				
MSMEG_4751	62				
MSMEG_4752	38				
MSMEG_4753	47				
MSMEG_4763	227				
MSMEG_4767	36				
MSMEG_4771	234				
MSMEG_4773	20				
MSMEG_4775	21				
MSMEG_4776	35				
MSMEG_4779	50				
MSMEG_4784	45				
MSMEG_4791	107				
MSMEG_4797	31				
MSMEG_4800	64				
MSMEG_4804	45				
MSMEG_4807	132				
MSMEG_4829	51				
MSMEG_4831	14				
MSMEG_4835	431				
MSMEG_4839	77				
MSMEG_4847	35				
MSMEG_4849	58				
MSMEG_4868	37				
MSMEG_4880	194				
MSMEG_4887	22				
MSMEG_4891	52				
MSMEG_4893	22				
MSMEG_4904	73				
MSMEG_4906	19				
MSMEG_4907	234				
MSMEG_4916	68				
MSMEG_4917	388				
MSMEG_4920	27				
MSMEG_4922	6				
MSMEG_4924	3				
MSMEG_4926	172				
MSMEG_4927	107				

MSMEG_4933	314				
MSMEG_4935	34				
MSMEG_4938	116				
MSMEG_4939	395				
MSMEG_4941	231				
MSMEG_4943	77				
MSMEG_4945	214				
MSMEG_4946	163				
MSMEG_4951	68				
MSMEG_4953	47				
MSMEG_4954	224				
MSMEG_4956	401				
MSMEG_4958	165				
MSMEG_4961	52				
MSMEG_4971	422				
MSMEG_4973	1				
MSMEG_4977	18				
MSMEG_4979	45				
MSMEG_4989	414				
MSMEG_4991	64				
MSMEG_4993	356				
MSMEG_4995	17				
MSMEG_4998	1				
MSMEG_5000	328				
MSMEG_5004	48				
MSMEG_5012	23				
MSMEG_5013	31				
MSMEG_5016	206				
MSMEG_5039	28				
MSMEG_5042	1				
MSMEG_5048	57				
MSMEG_5049	97				
MSMEG_5050	23				
MSMEG_5051	26				
MSMEG_5053	36				
MSMEG_5054	324				
MSMEG_5056	26				
MSMEG_5057	410				
MSMEG_5059	126				

MSMEG_5060	283				
MSMEG_5061	63				
MSMEG_5062	113				
MSMEG_5067	47				
MSMEG_5070	38				
MSMEG_5072	61				
MSMEG_5073	157				
MSMEG_5080	433				
MSMEG_5081	101				
MSMEG_5086	35				
MSMEG_5089	29				
MSMEG_5090	284				
MSMEG_5091	325				
MSMEG_5100	65				
MSMEG_5101	26				
MSMEG_5102	66				
MSMEG_5109	24				
MSMEG_5111	47				
MSMEG_5119	341				
MSMEG_5122	61				
MSMEG_5124	39				
MSMEG_5126	25				
MSMEG_5127	47				
MSMEG_5128	78				
MSMEG_5129	1				
MSMEG_5131	40				
MSMEG_5134	83				
MSMEG_5144	82				
MSMEG_5145	97				
MSMEG_5148	1				
MSMEG_5149	60				
MSMEG_5154	470				
MSMEG_5157	34				
MSMEG_5159	86				
MSMEG_5162	438				
MSMEG_5165	35				
MSMEG_5171	1				
MSMEG_5173	135				
MSMEG_5178	28				

MSMEG_5184	37				
MSMEG_5189	17				
MSMEG_5196	117				
MSMEG_5199	44				
MSMEG_5207	388				
MSMEG_5208	107				
MSMEG_5209	23				
MSMEG_5211	140				
MSMEG_5216	43				
MSMEG_5218	27				
MSMEG_5219	90				
MSMEG_5220	339				
MSMEG_5229	296				
MSMEG_5231	148				
MSMEG_5233	154				
MSMEG_5236	53				
MSMEG_5237	16				
MSMEG_5241	204				
MSMEG_5257	33				
MSMEG_5261	105				
MSMEG_5263	143				
MSMEG_5264	1				
MSMEG_5265	3				
MSMEG_5268	219				
MSMEG_5271	45				
MSMEG_5273	22				
MSMEG_5274	1				
MSMEG_5275	40				
MSMEG_5278	421				
MSMEG_5280	96				
MSMEG_5282	28				
MSMEG_5287	41				
MSMEG_5288	209				
MSMEG_5292	51				
MSMEG_5298	17				
MSMEG_5301	47				
MSMEG_5302	39				
MSMEG_5303	105				
MSMEG_5308	368				

MSMEG_5310	17				
MSMEG_5318	101				
MSMEG_5324	363				
MSMEG_5327	43				
MSMEG_5328	174				
MSMEG_5329	188				
MSMEG_5333	36				
MSMEG_5338	67				
MSMEG_5339	34				
MSMEG_5343	25				
MSMEG_5346	86				
MSMEG_5347	92				
MSMEG_5349	24				
MSMEG_5352	19				
MSMEG_5354	70				
MSMEG_5355	145				
MSMEG_5360	440				
MSMEG_5368	76				
MSMEG_5370	241				
MSMEG_5372	60				
MSMEG_5375	30				
MSMEG_5376	289				
MSMEG_5378	301				
MSMEG_5379	107				
MSMEG_5388	47				
MSMEG_5397	76				
MSMEG_5401	60				
MSMEG_5403	27				
MSMEG_5406	76				
MSMEG_5409	350				
MSMEG_5412	52				
MSMEG_5414	232				
MSMEG_5415	39				
MSMEG_5416	261				
MSMEG_5418	57				
MSMEG_5420	149				
MSMEG_5427	31				
MSMEG_5429	333				
MSMEG_5431	119				

MSMEG_5435	288				
MSMEG_5438	143				
MSMEG_5439	67				
MSMEG_5442	17				
MSMEG_5444	70				
MSMEG_5452	1				
MSMEG_5460	53				
MSMEG_5461	27				
MSMEG_5462	35				
MSMEG_5463	16				
MSMEG_5468	62				
MSMEG_5470	126				
MSMEG_5475	24				
MSMEG_5480	45				
MSMEG_5481	354				
MSMEG_5482	40				
MSMEG_5483	153				
MSMEG_5486	182				
MSMEG_5489	65				
MSMEG_5500	414				
MSMEG_5502	39				
MSMEG_5503	450				
MSMEG_5505	1				
MSMEG_5507	81				
MSMEG_5515	127				
MSMEG_5517	50				
MSMEG_5518	84				
MSMEG_5522	18				
MSMEG_5525	123				
MSMEG_5534	3				
MSMEG_5536	165				
MSMEG_5539	24				
MSMEG_5543	260				
MSMEG_5549	27				
MSMEG_5550	22				
MSMEG_5553	27				
MSMEG_5555	46				
MSMEG_5556	66				
MSMEG_5557	50				

MSMEG_5558	54				
MSMEG_5559	448				
MSMEG_5565	22				
MSMEG_5568	32				
MSMEG_5579	420				
MSMEG_5584	17				
MSMEG_5587	24				
MSMEG_5588	1				
MSMEG_5589	31				
MSMEG_5591	33				
MSMEG_5595	27				
MSMEG_5597	29				
MSMEG_5601	215				
MSMEG_5607	45				
MSMEG_5610	182				
MSMEG_5611	60				
MSMEG_5612	124				
MSMEG_5617	62				
MSMEG_5634	193				
MSMEG_5636	56				
MSMEG_5641	55				
MSMEG_5646	409				
MSMEG_5647	11				
MSMEG_5649	83				
MSMEG_5651	38				
MSMEG_5656	493				
MSMEG_5658	450				
MSMEG_5660	350				
MSMEG_5661	112				
MSMEG_5672	63				
MSMEG_5680	70				
MSMEG_5682	86				
MSMEG_5685	11				
MSMEG_5690	50				
MSMEG_5691	10				
MSMEG_5694	122				
MSMEG_5696	35				
MSMEG_5697	27				
MSMEG_5699	96				

MSMEG_5700	249				
MSMEG_5702	65				
MSMEG_5703	191				
MSMEG_5704	46				
MSMEG_5705	3				
MSMEG_5708	359				
MSMEG_5709	310				
MSMEG_5710	239				
MSMEG_5713	37				
MSMEG_5719	33				
MSMEG_5720	4				
MSMEG_5721	69				
MSMEG_5722	241				
MSMEG_5727	178				
MSMEG_5730	62				
MSMEG_5731	2				
MSMEG_5741	20				
MSMEG_5749	53				
MSMEG_5750	179				
MSMEG_5760	11				
MSMEG_5764	86				
MSMEG_5768	78				
MSMEG_5769	114				
MSMEG_5770	12				
MSMEG_5771	73				
MSMEG_5773	34				
MSMEG_5775	346				
MSMEG_5778	52				
MSMEG_5781	7				
MSMEG_5782	37				
MSMEG_5784	105				
MSMEG_5786	48				
MSMEG_5788	184				
MSMEG_5797	74				
MSMEG_5798	1				
MSMEG_5800	33				
MSMEG_5803	27				
MSMEG_5811	52				
MSMEG_5812	63				

MSMEG_5816	90				
MSMEG_5820	1				
MSMEG_5828	135				
MSMEG_5831	236				
MSMEG_5833	95				
MSMEG_5848	227				
MSMEG_5851	459				
MSMEG_5853	35				
MSMEG_5855	103				
MSMEG_5859	44				
MSMEG_5866	57				
MSMEG_5868	33				
MSMEG_5876	144				
MSMEG_5878	37				
MSMEG_5879	89				
MSMEG_5885	32				
MSMEG_5886	5				
MSMEG_5887	86				
MSMEG_5889	35				
MSMEG_5894	179				
MSMEG_5902	121				
MSMEG_5904	45				
MSMEG_5906	36				
MSMEG_5912	35				
MSMEG_5916	81				
MSMEG_5920	241				
MSMEG_5925	23				
MSMEG_5929	44				
MSMEG_5933	59				
MSMEG_5934	30				
MSMEG_5937	89				
MSMEG_5939	13				
MSMEG_5940	157				
MSMEG_5943	130				
MSMEG_5944	104				
MSMEG_5952	474				
MSMEG_5956	144				
MSMEG_5959	399				
MSMEG_5962	270				

MSMEG_5967	114				
MSMEG_5968	32				
MSMEG_5969	224				
MSMEG_5970	29				
MSMEG_5971	229				
MSMEG_5976	78				
MSMEG_5981	242				
MSMEG_5988	121				
MSMEG_5989	30				
MSMEG_5995	32				
MSMEG_5997	182				
MSMEG_6001	43				
MSMEG_6006	158				
MSMEG_6010	126				
MSMEG_6013	86				
MSMEG_6018	50				
MSMEG_6019	250				
MSMEG_6020	160				
MSMEG_6021	46				
MSMEG_6022	37				
MSMEG_6024	207				
MSMEG_6027	77				
MSMEG_6029	2				
MSMEG_6031	31				
MSMEG_6032	74				
MSMEG_6033	36				
MSMEG_6038	31				
MSMEG_6039	161				
MSMEG_6040	22				
MSMEG_6042	198				
MSMEG_6044	2				
MSMEG_6056	192				
MSMEG_6064	152				
MSMEG_6068	24				
MSMEG_6073	66				
MSMEG_6076	257				
MSMEG_6077	20				
MSMEG_6078	31				
MSMEG_6080	25				

MSMEG_6083	62				
MSMEG_6085	17				
MSMEG_6090	120				
MSMEG_6091	75				
MSMEG_6092	59				
MSMEG_6098	66				
MSMEG_6104	115				
MSMEG_6105	75				
MSMEG_6112	305				
MSMEG_6113	45				
MSMEG_6114	70				
MSMEG_6131	66				
MSMEG_6136	130				
MSMEG_6141	22				
MSMEG_6145	135				
MSMEG_6148	354				
MSMEG_6149	107				
MSMEG_6153	1				
MSMEG_6156	107				
MSMEG_6157	44				
MSMEG_6159	183				
MSMEG_6160	46				
MSMEG_6162	107				
MSMEG_6171	423				
MSMEG_6180	68				
MSMEG_6182	31				
MSMEG_6185	87				
MSMEG_6198	23				
MSMEG_6199	117				
MSMEG_6201	392				
MSMEG_6219	112				
MSMEG_6221	222				
MSMEG_6224	27				
MSMEG_6227	39				
MSMEG_6229	35				
MSMEG_6237	236				
MSMEG_6239	42				
MSMEG_6240	155				
MSMEG_6241	203				

MSMEG_6242	41				
MSMEG_6243	34				
MSMEG_6245	144				
MSMEG_6250	42				
MSMEG_6254	117				
MSMEG_6257	41				
MSMEG_6259	40				
MSMEG_6266	73				
MSMEG_6270	327				
MSMEG_6271	117				
MSMEG_6280	117				
MSMEG_6282	71				
MSMEG_6284	18				
MSMEG_6286	146				
MSMEG_6291	55				
MSMEG_6300	28				
MSMEG_6302	1				
MSMEG_6303	31				
MSMEG_6314	351				
MSMEG_6316	47				
MSMEG_6334	291				
MSMEG_6340	46				
MSMEG_6344	34				
MSMEG_6346	34				
MSMEG_6348	44				
MSMEG_6352	107				
MSMEG_6353	24				
MSMEG_6354	488				
MSMEG_6356	126				
MSMEG_6358	325				
MSMEG_6363	106				
MSMEG_6364	113				
MSMEG_6365	99				
MSMEG_6369	72				
MSMEG_6379	376				
MSMEG_6380	258				
MSMEG_6381	22				
MSMEG_6383	203				
MSMEG_6385	29				

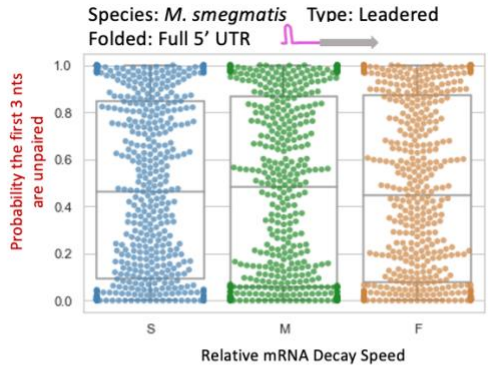
MSMEG_6386	48				
MSMEG_6390	38				
MSMEG_6391	68				
MSMEG_6393	91				
MSMEG_6394	151				
MSMEG_6398	73				
MSMEG_6399	68				
MSMEG_6401	62				
MSMEG_6404	1				
MSMEG_6405	103				
MSMEG_6408	307				
MSMEG_6410	31				
MSMEG_6411	303				
MSMEG_6414	36				
MSMEG_6416	110				
MSMEG_6419	30				
MSMEG_6426	81				
MSMEG_6427	64				
MSMEG_6428	53				
MSMEG_6430	142				
MSMEG_6432	42				
MSMEG_6433	45				
MSMEG_6434	56				
MSMEG_6438	62				
MSMEG_6446	16				
MSMEG_6447	364				
MSMEG_6448	11				
MSMEG_6451	35				
MSMEG_6452	36				
MSMEG_6456	38				
MSMEG_6457	41				
MSMEG_6477	25				
MSMEG_6480	29				
MSMEG_6481	307				
MSMEG_6484	77				
MSMEG_6487	160				
MSMEG_6497	83				
MSMEG_6498	42				
MSMEG_6508	204				

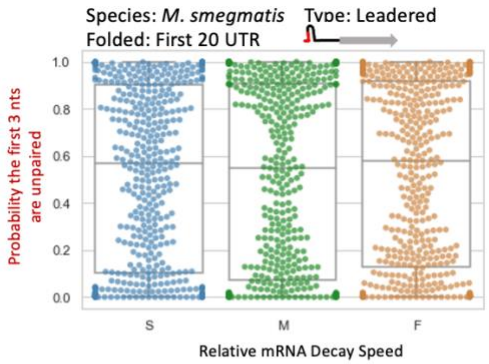
MSMEG_6511	24				
MSMEG_6515	1				
MSMEG_6518	132				
MSMEG_6519	324				
MSMEG_6524	119				
MSMEG_6525	1				
MSMEG_6528	124				
MSMEG_6535	48				
MSMEG_6538	151				
MSMEG_6540	42				
MSMEG_6541	109				
MSMEG_6543	90				
MSMEG_6544	185				
MSMEG_6552	47				
MSMEG_6555	66				
MSMEG_6563	52				
MSMEG_6566	63				
MSMEG_6568	70				
MSMEG_6577	120				
MSMEG_6579	52				
MSMEG_6581	67				
MSMEG_6587	47				
MSMEG_6588	42				
MSMEG_6590	1				
MSMEG_6596	25				
MSMEG_6599	42				
MSMEG_6601	1				
MSMEG_6603	35				
MSMEG_6604	27				
MSMEG_6618	3				
MSMEG_6619	168				
MSMEG_6630	29				
MSMEG_6634	301				
MSMEG_6635	23				
MSMEG_6636	40				
MSMEG_6637	65				
MSMEG_6638	339				
MSMEG_6640	480				
MSMEG_6642	280				

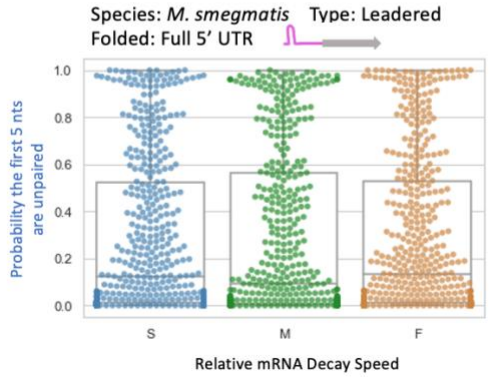
MSMEG_6649	323				
MSMEG_6650	180				
MSMEG_6653	29				
MSMEG_6654	111				
MSMEG_6660	39				
MSMEG_6663	56				
MSMEG_6678	34				
MSMEG_6686	65				
MSMEG_6690	96				
MSMEG_6698	182				
MSMEG_6700	73				
MSMEG_6726	105				
MSMEG_6727	47				
MSMEG_6730	37				
MSMEG_6739	94				
MSMEG_6740	59				
MSMEG_6743	37				
MSMEG_6744	115				
MSMEG_6747	44				
MSMEG_6754	43				
MSMEG_6758	54				
MSMEG_6759	471				
MSMEG_6760	343				
MSMEG_6761	65				
MSMEG_6764	22				
MSMEG_6771	107				
MSMEG_6772	53				
MSMEG_6781	371				
MSMEG_6783	23				
MSMEG_6786	441				
MSMEG_6787	124				
MSMEG_6789	88				
MSMEG_6800	124				
MSMEG_6802	91				
MSMEG_6804	44				
MSMEG_6808	35				
MSMEG_6821	58				
MSMEG_6822	58				
MSMEG_6826	32				

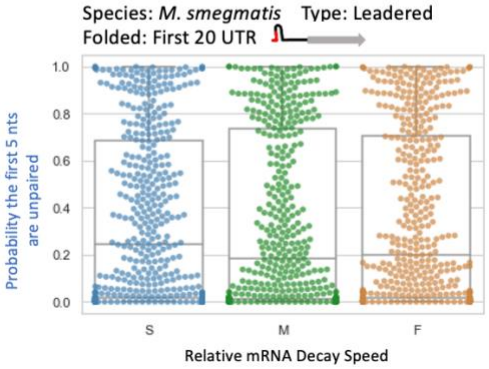
MSMEG_6843	256				
MSMEG_6859	23				
MSMEG_6861	70				
MSMEG_6864	1				
MSMEG_6865	118				
MSMEG_6874	64				
MSMEG_6881	37				
MSMEG_6884	56				
MSMEG_6891	425				
MSMEG_6892	263				
MSMEG_6896	78				
MSMEG_6897	93				
MSMEG_6900	214				
MSMEG_6903	88				
MSMEG_6911	443				
MSMEG_6914	1				
MSMEG_6919	50				
MSMEG_6922	38				
MSMEG_6925	244				
MSMEG_6932	1				
MSMEG_6933	37				
MSMEG_6934	18				
MSMEG_6935	45				
MSMEG_6936	42				
MSMEG_6941	337				
MSMEG_6942	116				
MSMEG_6946	168				
MSMEG_6947	226				

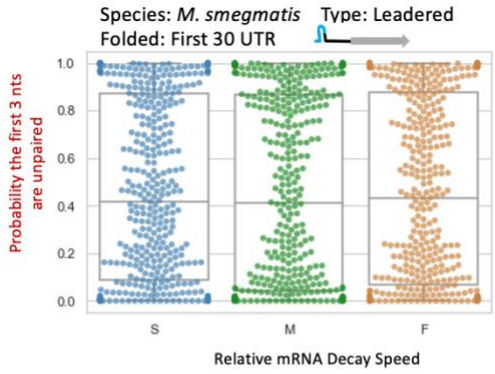
Appendix B.

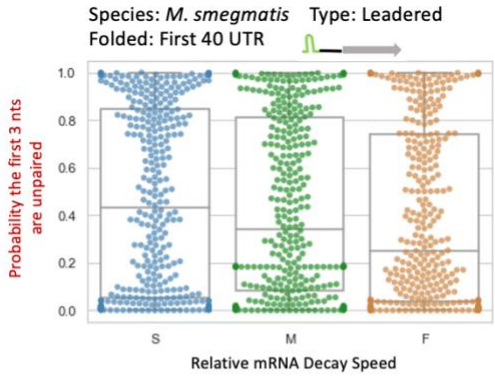

Species: M. smegmatis Type: Leadered
Folded: Full 5' UTR
Probability the first 3 nts are unpaired
1.0
0.8
0.6
0.4
0.2
0.0
S
M
F
Relative mRNA Decay Speed

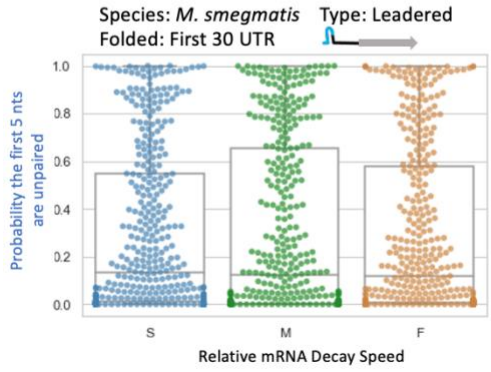

Species: M. smegmatis Type: Leadered
Folded: First 20 UTR
Probability the first 3 nts are unpaired
1.0
0.8
0.6
0.4
0.2
0.0
S
M
F
Relative mRNA Decay Speed

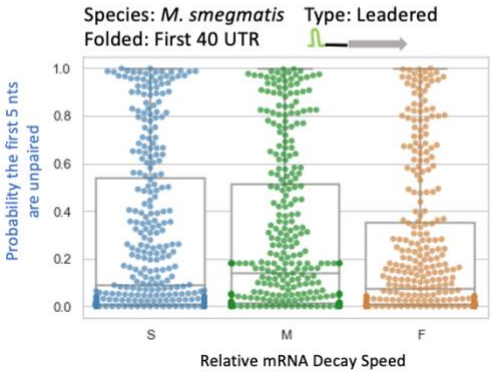

Species: M. smegmatis Type: Leadered
Folded: Full 5' UTR
Probability the first 5 nts are unpaired
1.0
0.8
0.6
0.4
0.2
0.0
S
M
F
Relative mRNA Decay Speed

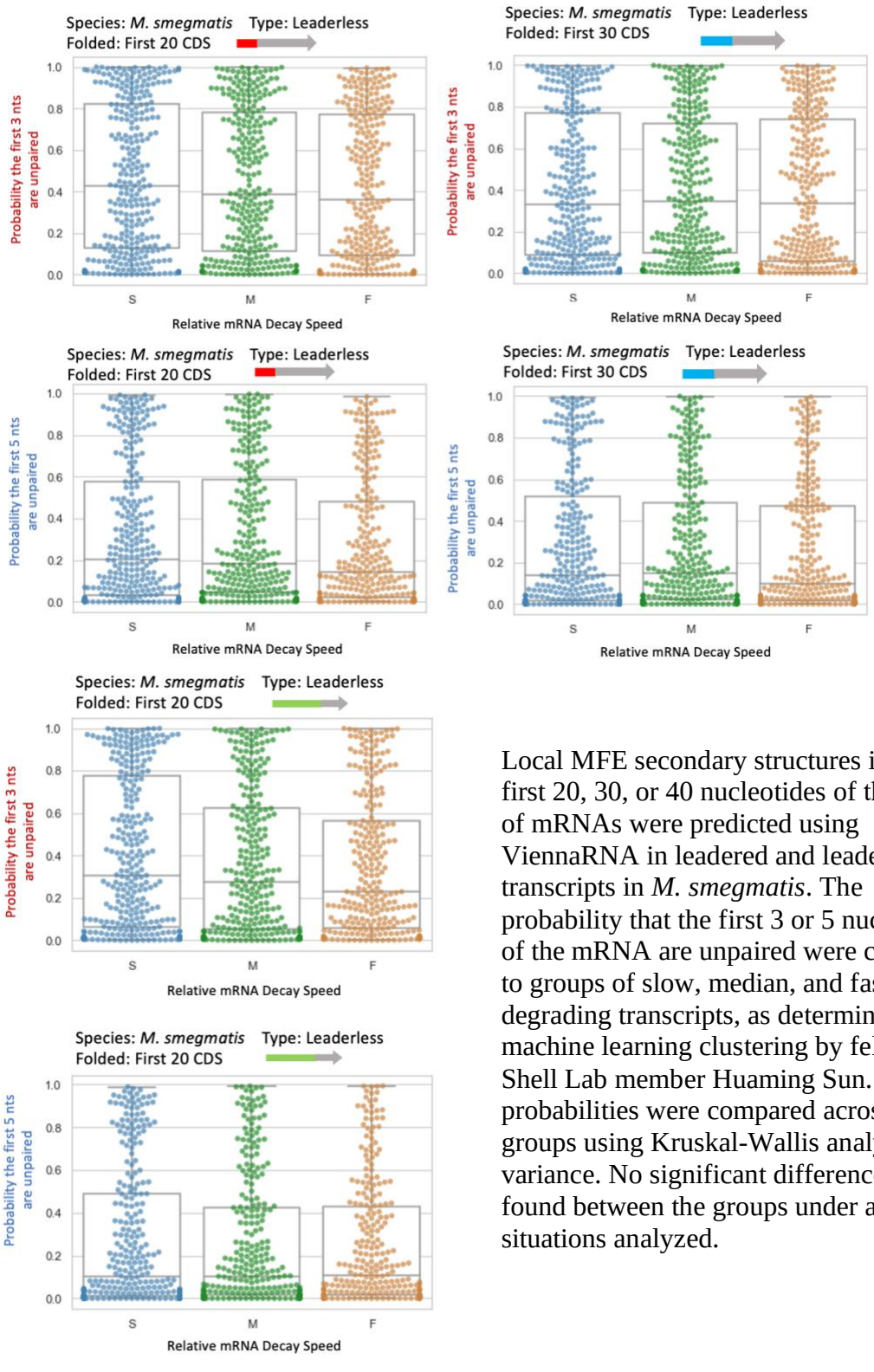

Species: M. smegmatis Type: Leadered
Folded: First 20 UTR
Probability the first 5 nts are unpaired
1.0
0.8
0.6
0.4
0.2
0.0
S
M
F
Relative mRNA Decay Speed


Species: M. smegmatis Type: Leadered
Folded: First 30 UTR
Probability the first 3 nts are unpaired
1.0
0.8
0.6
0.4
0.2
0.0
S
M
F
Relative mRNA Decay Speed


Species: M. smegmatis Type: Leadered
Folded: First 40 UTR
Probability the first 3 nts are unpaired
1.0
0.8
0.6
0.4
0.2
0.0
S
M
F
Relative mRNA Decay Speed


Species: M. smegmatis Type: Leadered
Folded: First 30 UTR
Probability the first 5 nts are unpaired
1.0
0.8
0.6
0.4
0.2
0.0
S
M
F
Relative mRNA Decay Speed


Species: M. smegmatis Type: Leadered
Folded: First 40 UTR
Probability the first 5 nts are unpaired
1.0
0.8
0.6
0.4
0.2
0.0
S
M
F
Relative mRNA Decay Speed



Local MFE secondary structures in the first 20, 30, or 40 nucleotides of the 5' end of mRNAs were predicted using ViennaRNA in leadered and leaderless transcripts in *M. smegmatis*. The probability that the first 3 or 5 nucleotides of the mRNA are unpaired were compared to groups of slow, median, and fast degrading transcripts, as determined by machine learning clustering by fellow Shell Lab member Huaming Sun. Median probabilities were compared across the groups using Kruskal-Wallis analysis of variance. No significant difference was found between the groups under any of the situations analyzed.