RNA BIOLOGY
2016, VOL. 0, NO. 0, 1–12
http://dx.doi.org/10.1080/15476286.2016.1144010

Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER

Novel small RNA (sRNA) landscape of the starvation-stress response transcriptome of *Salmonella enterica* serovar typhimurium

Shivam V. Amin[a,#], Justin T. Roberts[a,#], Dillon G. Patterson[a], Alexander B. Coley[a], Jonathan A. Allred[a], Jason M. Denner[a], Justin P. Johnson[a], Genevieve E. Mullen[a], Trenton K. O'Neal[a], Jason T. Smith[a], Sara E. Cardin[a], Hank T. Carr[a], Stacie L. Carr[a], Holly E. Cowart[a], David H. DaCosta[a], Brendon R. Herring[a], Valeria M. King[a], Caroline J. Polska[a], Erin E. Ward[a], Alice A. Wise[a], Kathleen N. McAllister[b], David Chevalier[c], Michael P. Spector[b], and Glen M. Borchert[a,d]

[a]Department of Biology; University of South Alabama; Mobile, AL ; [b]Department of Biomedical Sciences; University of South Alabama; Mobile, AL ; [c]Department of Biology; East Georgia State College; Swainsboro, GA ; [d]Department of Pharmacology; USA College of Medicine; Mobile, AL

ABSTRACT

Small RNAs (sRNAs) are short (~50–200 nucleotides) noncoding RNAs that regulate cellular activities across bacteria. *Salmonella enterica* starved of a carbon-energy (C) source experience a host of genetic and physiological changes broadly referred to as the starvation-stress response (SSR). In an attempt to identify novel sRNAs contributing to SSR control, we grew log-phase, 5-h C-starved and 24-h C-starved cultures of the virulent *Salmonella enterica* subspecies enterica serovar Typhimurium strain SL1344 and comprehensively sequenced their small RNA transcriptomes. Strikingly, after employing a novel strategy for sRNA discovery based on identifying dynamic transcripts arising from "gene-empty" regions, we identify 58 wholly undescribed *Salmonella* sRNA genes potentially regulating SSR averaging an ~1,000-fold change in expression between log-phase and C-starved cells. Importantly, the expressions of individual sRNA loci were confirmed by both comprehensive transcriptome analyses and northern blotting of select candidates. Of note, we find 43 candidate sRNAs share significant sequence identity to characterized sRNAs in other bacteria, and ~70% of our sRNAs likely assume characteristic sRNA structural conformations. In addition, we find 53 of our 58 candidate sRNAs either overlap neighboring mRNA loci or share significant sequence complementarity to mRNAs transcribed elsewhere in the SL1344 genome strongly suggesting they regulate the expression of transcripts via antisense base-pairing. Finally, in addition to this work resulting in the identification of 58 entirely novel *Salmonella enterica* genes likely participating in the SSR, we also find evidence suggesting that sRNAs are significantly more prevalent than currently appreciated and that *Salmonella* sRNAs may actually number in the thousands.

ARTICLE HISTORY
Received 9 December 2015
Revised 13 January 2016
Accepted 14 January 2016

KEYWORDS
carbon-starvation; noncoding RNA; *Salmonella*; sRNA; starvation-stress; transcriptomics

Introduction

Salmonella enterica serovars are rod-shaped, Gram-negative bacterium implicated in food-borne diarrheal and systemic disease, causing thousands of human deaths worldwide each year.[1,2] This bacterium is a facultative anaerobe and chemoorganotroph capable of acquiring both carbon and electrons from D-glucose and other sugars. After ingestion by a new host, most *Salmonella* serovars invade the mucosal epithelium of the small intestine to produce inflammation and diarrheal disease. Some serovars can spread beyond the intestinal immune tissue causing systemic infections (*e.g.*, bacteremia and typhoid fever) and can also set up chronic infections within the gall bladder and may survive within macrophages long after symptoms of the primary infection disappear.[3]

Like many species of bacteria that cause human disease, a common stress *Salmonella* face is the starvation of essential nutrients, such as a carbon-energy (C) source.[4] In particular, *S. enterica* starved for a C-source experience a host of genetic and physiological changes referred to as the starvation-stress response (SSR). These C-starvation-induced changes result in a metabolically "reprogrammed" *Salmonella* cell able to: more efficiently utilize any nutrients that become available in its environment, resist the effects of long-term C-starvation, and resist multiple other environmental stresses.[4] These metabolic changes affect the production of proteins/enzymes, RNAs, lipids, and a wide-array of other cellular machinery components.[5] The expression of several genes can exhibit ≥10 -fold induction in C-starved cells relative to that of non-starved cells. Interestingly, ~20% of the up-regulated genes expressed during the SSR have unknown or only putative functions with several being unique to certain *Salmonella* serovars.[4] Currently, the SSR is found to be globally positively or negatively regulated by specific promoter-binding sigma factors (e.g.,, σ^S and σ^E),

CONTACT Glen M. Borchert borchert@southalabama.edu, Michael P. Spector mspector@southalabama.edu

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/krnb.

Supplemental material data for this article can be accessed on the publisher's website.

[#] These authors contributed equally to this work.

secondary messengers (*e.g.*, cAMP and ppGpp), and 2-component sensor-kinase/response regulator systems.[4,6,7] However, there are still many unknowns concerning the processes through which *Salmonella* serovars respond to and survive carbon-starvation. As hundreds of novel small RNAs (sRNAs) are just now being described, and since each of these represents a unique, undescribed genetic regulator, sRNAs may well provide the necessary link to bridge some longstanding gaps in knowledge.

Small RNAs (sRNAs) are short noncoding RNA sequences (50–200 base pairs) that regulate many cellular activities by binding to complementary target sequences within mRNAs (mRNAs) to affect their translation.[8,9] These molecules are found in numerous species across all domains of life.[10,11] Until the turn of this century, sRNAs were largely an unknown and unexplored area of research, but they have recently been shown to modulate numerous physiological mechanisms and pathways in a number of different organisms, from bacteria to humans.[9]

While many sRNAs inhibit translation, others inhibit the mobility of the transposon from which they originated, some inhibit genomic locus transcription, and some can even exhibit positive regulation of mRNA expressions through an uncharacterized mechanism(s).[9,12-15] That said, in most bacteria, sRNAs are transcribed as small, distinct transcripts containing nucleotide sequences that are either perfectly or imperfectly complementary to, and bind to, target sequences in one or more mRNAs, thereby regulating the translation or degradation of the target mRNA (s).[16,17]. Antisense sRNAs can be divided into 2 categories based on their location relative to their target gene. When the sRNA gene and its target gene partially overlap, the sRNA is referred to as a *cis*-sRNA. It is encoded on the DNA strand opposite to the one encoding its target mRNA gene, and the resultant sRNA and mRNA transcripts complement one another perfectly over a short, finite region.[10] Researchers know little about the regulatory function(s) of *cis*-encoded antisense sRNAs, but it has been hypothesized that some are expressed constitutively[9] In contrast, *trans*-encoded sRNAs are located at distinct sites on the chromosome from the genes that they regulate and are typically found in the intergenic regions of bacterial genomes[8] In addition, they do not require perfect complementarity with their mRNA targets.[10] After binding to a mRNA transcript, a complementary *trans*-encoded sRNA typically labels the target mRNA for degradation by RNAses or directly inhibits translation by preventing ribosome binding.[17,18] Most of the currently identified *trans*-encoded sRNAs in bacterial species bind to the 5′ untranslated region (UTR) of mRNAs and prevent ribosomal binding.[17] Target sites, or sequences where a *trans*-encoded sRNA imperfectly complements and binds with an mRNA, are generally 10–25 bp long.[9]

Bioinformatic prediction of sRNAs remains an inexact science, and determination of the function(s) of particular sRNAs is an even more daunting task.[10] That said, the advent of RNA-Seq and associated bioinformatic tools has now made the analysis of bacterial transcriptomic data considerably more extensive and efficient. Since it has been well established that the SSR of *S.* Typhimurium cells elicited by C-starvation results in global changes in metabolic pathways, many of which manifest in obvious phenotypic differences in relation to non-starved cells,[4] we hypothesized that C-starved *Salmonella* might express unique sRNAs not observed in non-starved *Salmonella*, and non-starved *Salmonella* might likewise express sRNAs not transcribed under C-starvation conditions. Therefore, in this study we have attempted to identify and characterize novel sRNAs that are differentially expressed during the SSR by conducting RNA-Seq of the *S.* Typhimurium strain SL1344 transcriptome grown under non-starvation and C-starvation conditions.

Results

S. Typhimurium differentially expresses sRNAs under different growth conditions

To determine whether *S.* Typhimurium differentially expresses novel sRNAs in response to carbon-source starvation conditions, (non-starved) log-phase, 5-h and 24-h C-starved cell populations were generated, RNA was isolated, from the 3 cultures, and sequenced using a small RNA RNA-Seq protocol resulting in roughly 6 million paired-end 100 bp reads for each culture (Table 1). As sRNAs typically originate from noncoding regions of the genome, we began by identifying all chromosomal intergenic regions longer than 50 bps from the SL1344 chromosome employing a methodology similar to previous reports.[19,20] In all, we identified 3,035 chromosomal locations meeting this criteria (Fig. 1) then focused our search for novel sRNAs on these specific regions.

Next, following the identification of these *S.* Typhimurium gene empty locations, RNA-Seq reads from our 3 cell populations were uniquely aligned to these intergenic regions using BLAST+. Strikingly, we found 2,620 of these regions significantly aligned to at least one of our unique transcriptome reads. As such, in order to limit our focus to those sRNAs that are differentially expressed between the 3 growth conditions, the total number of reads aligning to each intergenic region was calculated for each condition and regions exhibiting significant, dynamic changes in transcription were selected resulting in the identification of 63 regions of interest (Fig. 2, Tables 2 and 3). Reads aligning to these regions were then extracted and consolidated into a consensus sequence representing a putative sRNA. In all, we identified 63 candidate sRNAs that warranted further investigation. Of note, although we describe each of our novel sRNAs as intergenic, in actuality 21 of these novel sRNAs do partially overlap the 5′ or 3′ ends of known protein coding gene loci.

Table 1. Number of raw, single-reads generated from RNA-seq. BLAST+ alignments between raw reads and intergenic regions.

Sample	Number of Raw Reads	Total Number of Alignments[a]
Log-phase R1	3,228,873	353,158
Log-phase R2	3,266,700	348,016
5-h C-starved R1	3,557,939	197,336
5-h C-starved R2	3,722,680	195,037
24-h C-starved R1	2,601,209	198,554
24-h C-starved R2	2,661,468	197,128

[a]≥40 bp and ≥98 % identity

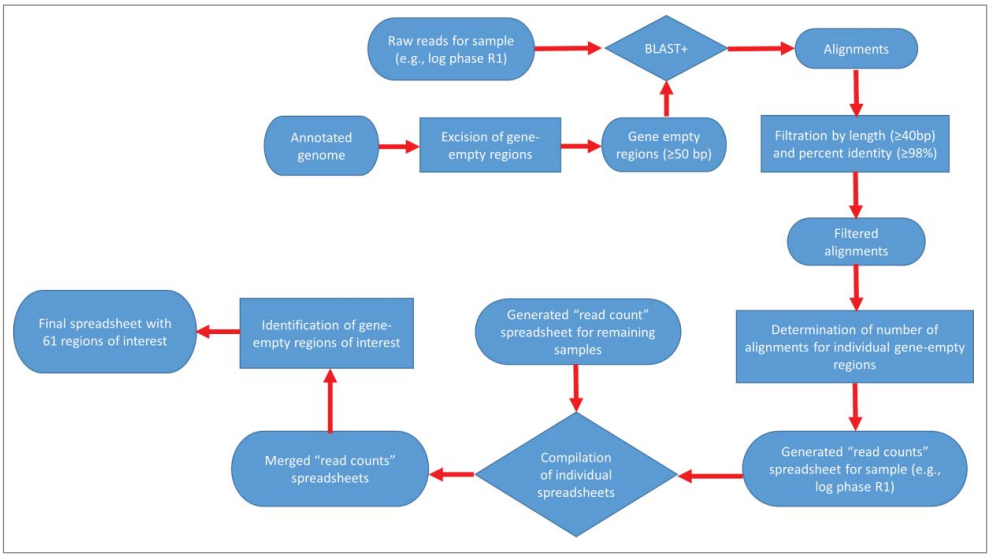

Figure 1. Workflow for identifying candidate sRNAs. Reads belonging to a particular sample (e.g., log-phase R1) were aligned to excised intergenic regions of the SL1344 chromosome. "Read counts" file of a given sample showed the number of alignments to each individual intergenic region. Compilation of "read counts" files allowed for comparison of expression levels across all 6 samples for a specific intergenic region. 63 regions of interest were identified.

These overlapping sRNAs would still predominately be considered intergenic as they average 76.6 bps in length with an average protein coding gene overlap of 26.7 bps.

58 of 63 candidate sRNAs are novel

Importantly, 2 major transcriptome analyses of SL1344 identifying significant sets of *S.* Typhimurium sRNAs have previously been reported. The first report identifying large numbers of sRNAs in SL1344 was published by Kröger et al.[2] in 2012 detailing ~140 novel sRNAs expressed under a single growth condition. A subsequent study by the same group[21] extended the SL1344 transcriptomic landscape by searching for sRNAs regulated by infection-relevant conditions identifying another ~140 novel sRNAs essentially doubling the number of characterized *S.* Typhimurium sRNAs (280). In order to determine if our 63 candidates overlapped with these and had previously been reported, we examined each of our candidate regions in the SalCom database (http://bioinf.gen.tcd.ie/cgi-bin/salcom.pl?_HL). To our surprise, we found only 5 of our 63 candidate sRNAs (Table 3) had been previously identified – sRNA176086,[22] sRNA9247442, sRNA17846792, sRNA2073265,[21] and sRNA4236850[23] – via comparative analyses with the *E. coli* chromosome,[22] or as being expressed in early stationary-phase cultures grown in nutrient-rich medium2, or as being regulated by infection-relevant conditions[21] or as being associated with the Hfq RNA binding protein,[23] respectively. Importantly, this indicates that 58 of the 63 sRNAs identified under the conditions tested here have not been previously reported in *Salmonella* or other *Enterobacteriaceae*.

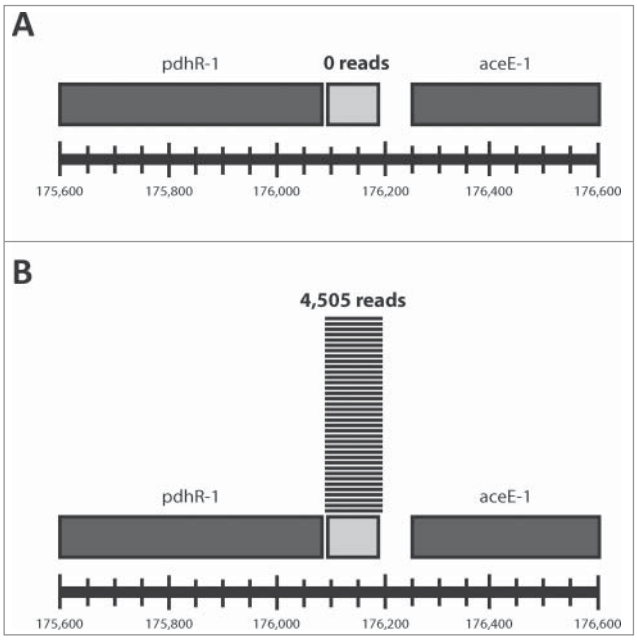

Figure 2. Illustration depicting sequence assembly as visualized with Tablet. (**A**) and (**B**) represent 24-h C-starved and log phase cultures, respectively. Putative sRNA is illustrated as a light gray box and neighboring protein coding genes as dark gray. Sequence reads mapping to the putative sRNA are indicated as horizontal lines above the region with total number indicated. Like all of our candidate sRNAs, this sRNA maps to an intergenic region of the SL1344 chromosome. This particular sRNA begins at position 176086 in the SL1344 chromosome and is approximately 90 nts long.

Table 2. Intergenic region expression data. Region of interest spanned 385 bps from position 4,217,226 to position 4,217,612 in the SL1344 chromosome.

Intergenic region of interest	Sample	Number of Alignments[a]
sRNA 4217226_4217612	Log-phase R1	21,169
	Log-phase R2	20,763
sRNA 4217226_4217612	5-h C-starved R1	3,293
	5-h C-starved R2	3,247
sRNA 4217226_4217612	24-h C-starved R1	80
	24-h C-starved R2	76

[a]≥40 bp and ≥98 % identity

Table 3. Summary of putative sRNAs in actively dividing and SSR *Salmonella enterica*. This table lists the following characteristics of all candidate sRNAs: strand (plus / minus / both), known sRNA similarity, structure (HS/HP), putative targets and expression patterns log-phase, 5-h C-starved and 24-h C-starved cells.

sRNA	Strand[a]	Known sRNA Similarity	Structure[b]	xPutative Gene Targets[Gene Products] / (Overlap)[c]	Exp. In Log[d]	Exp. In 5 -h C-st[d]	Exp. In 24-h C-st[d]
sRNA176086	Plus	–	HS	• *cis-pdhR* [pyruvate dehydrogenase (Pyr DH) complex repressor] (3′)*cis-aceE* [Pyr DH decarboxylase subunit E1] (5′)	4,505	621	0
sRNA201302	Both	**E. coli* uncharacterized sRNA	HS	• *cis-hpt* (hypoxanthine phosphoribosyltransferase] (3′)*cis-yadF* [carbonic anhydrase] (3′)	230	721	0
sRNA294324	Plus	*S. enterica* uncharacterized sRNA	HS	• *cis-rrlH* [23S rRNA] (3′)*cis-rrfH* [5S rRNA] (5′)	16,711	3,899	0
sRNA523910	Minus	Bacterial small SRP	HP	• *cis-ybaY1* (5′), *cis*-4.5S RNA(5′)*cis- ybaZ1* (3′)	1,060	1,975	9,640
sRNA570517	Both	–	HS	• *cis*-STM0509 [OmpC-like outer membrane protein] (3′) *cis-sfbA* [Fe^{3+}- / pH-inducible ABC transporter substrate-binding protein] (5′)	798	0	0
sRNA697722	Minus	–	HS	• *cis*-STM0636 [hypothetical protein] (5′)*cis-dacA* [PBP5;D-alanyl-D-alanine carboxypeptidase] (3′)	79	140	635
sRNA924744	Both	–	HS	• *cis-bssR* [biofilm formation regulatory protein] (3′)	115	320	0
sRNA1130547	Minus	–	HS	• *cis-yccK* [sulfurtransferase TusE] (5′)*cis*-STM1085 [YccA-like Inner membrane protein; possible modulator of FtsH protease] (3′)	42	122	5,044
sRNA1170414	Plus	–	HS	• *cis-phoH* [phosphate starvation-inducible ATP-binding protein PsiH] (5′)	53	383	1105
sRNA1186573	Minus	–	HS	• *cis-csgD* [transcriptional activator protein of curli pili production] (5′)	0	200	0
sRNA1212673	Both	*E. coli* uncharacterized sRNA	HS	• *cis-mviN* (STM1170) [putative MurJ peptidoglycan lipid II flippase] (3′)*cis-flgN* [FlgK/FlgL export chaperone] (3′)	1,147	51	28
sRNA1253515	Plus	Bacterial RNaseP	HS	• *cis-ycfR* [hypothetical outer membrane / secreted protein] (5′)	0	0	212
sRNA1288534	Plus	*Campylobacter jejuni* uncharacterized sRNA	HS	–	0	548	71
sRNA1289188		–	HS	–	0	548	71
sRNA1292328	Plus	–	HS	• *cis*-STM1250 [cytoplasmic protein] (3′)*cis*-STM1251 [small heat shock IbpA/B-like molecular chaperone] (5′)	0	0	429
sRNA1305985	Both	*E. coli* RtT	HS	• *trans*-SL1344_3673A [pseudogene; *tdh-yibD* intergenic region] *cis*-STM1266 [MerR-type transcriptional regulator] (3′)*cis*-STM1267 [YmgB-superfamily cytoplasmic protein] (3′)	111	1,720	1,440
sRNA1318601	Minus	*E. coli* uncharacterized sRNA	HS	• *cis*-STM1284 [metal ion-dependent adhesion site (MIDAS)-containing protein] (5′)*cis-yeaG* [PrkA-family serine protein kinase] (3′)	23	270	576
sRNA1326799	Plus	*Hyperthermus butylicus* uncharacterized sRNA	HS	• *cis-yeaA(msrB)* [peptide methionine sulfoxide reductase – oxidative stress protection] (5′)	34	114	43
sRNA1409050	Plus	–	HS	• *cis-sufA* [iron-sulfur cluster assembly scaffold protein] (5′)	0	253	600
sRNA1482545	Both	–	HS	• *cis-pdxY* [pyridoxamine / pyridoxal kinase] (3′)*cis-gst* [glutathionine S-transferase – detoxifiying enzyme] (3′)	0	0	143
sRNA1548865	Plus	Bacterial T-box	HS	• *cis*-STM1513 [putative general stress protein; GsiB] (5′)	0	221	0
sRNA1784679	Both	**Pseudomonas aeruginosa* PrrF2	HP	• *cis*-STM1731 [KatN-like (Mn^{2+})-catalase] (3′)*cis-ompW* [outer membrane (porin) protein W] (3′)	0	473	0
sRNA1799792	Plus	–	HS	• *cis-adhE* [bifunctional acetaldehyde-CoA / alcohol dehydrogenase] (5′)	3,307	793	128
sRNA1799950	Plus	–	HS	• *cis-adhE* [bifunctional acetaldehyde-CoA / alcohol dehydrogenase] (5′)	3,307	793	128
sRNA1809591	Plus	*E. coli* RtT	HS	• *cis-tyrT* [tRNA-Tyr] (3′)*cis-tyrV* [tRNA-Tyr] (5′)	1,535	313	238
sRNA1851877	Plus	–	HS	• *cis-treA* [periplasmic trehalase] (5′)• *cis*-STM1795 [GdhA-family glutamate dehydrogenase] (3′)	0	279	373
sRNA1888744	Plus	*Hyperthermus butylicus* uncharacterized sRNA	HS	• *cis-manZ* [mannose-specific PTS Enzyme IID component] (3′)• *cis*-STM1833 [putative membrane protein] (5′)	0	0	282
sRNA1897579	Minus	*E. coli* SsrA	HS	• *cis-htpX* [heat-shock Zn^{2+}-dependent protease / chaperone] (CBW17867)(3′)	0	0	291

(*Continue on next page*)

Table 3. (*Continued*)

sRNA	Strand[a]	Known sRNA Similarity	Structure[b]	xPutative Gene Targets[Gene Products] / (Overlap)[c]	Exp. In Log[d]	Exp. In 5 -h C-st[d]	Exp. In 24-h C-st[d]
sRNA1932297	Both	–	HS	• *cis-yebE* [putative metal-binding inner membrane protein (tellurium resistance TerB-like] (5′)*cis*-STM1881 [YebF-like (putative colicin immunity) protein] (3′)	332	142	1,469
sRNA2061316	Minus	–	HS	• *cis*-SL1344_1952 [predicted bacteriophage protein / endonuclease] (5′)*cis*-SL1344_1953 [predicted bacteriophage protein] (3′)	0	27	400
sRNA2073265	Plus	–	HS	• *cis*-SL1344_1968 [predicted bacteriophage protein] (5′)	2,134	731	38
sRNA2176581	Both	RNaseP in multiple organisms	HP	• *cis-rfbB* [dTDP-glucose-4,6-dehydratase] (5′)*cis-galF* [UTP–glucose-1-phosphate uridylyltransferase subunit] (3′)	752	1,830	60
sRNA2285422	Both	–	HS	• *cis-mglB* [galactose / methyl-galactoside ABC transporter substrate-binding protein] (3′)	0	296	77
sRNA2296504	Minus	**E. coli* Leu Leader	HP	• *cis-lysP* [APC family lysine-specific permease] (3′)*trans-cirA* [(TonB-dependent) catecholate siderophore receptor] (5′)	1,013	255	76
sRNA2416394	Both	–	HS	• *cis-menF* [(menaquinone-specific) isochorismate synthase] (5′)*cis-elaB* [conserved putative inner membrane (ribosome-binding) protein] (3′)	40	430	0
sRNA2576882	Both	**Streptococcus suis* uncharacterized sRNA	HS	• *cis-tnpA_1b* [insertion sequence IS200 transposase] (3′)	0	0	233
sRNA2594511	Both	–	HS	• *cis-acrD* [RND family aminoglycoside / multidrug efflux pump] (3′)*cis-ypfM* [stress-induced small enterobacterial protein in *E. coli*; annotated as STnc250] (3′)	27	31	1,232
sRNA2759412	Minus	–	HS	• *cis-gltW* [tRNA-Glu] (5′)	157	1,935	67
sRNA2802132	Minus	–	HP	• *cis-rrsG* [16S rRNA] (5′)	11,628	1,067	0
sRNA2838164	Both	–	HS	• *cis-smpB* [SsrA (tmRNA)-binding protein] (3′)	47	227	3,944
sRNA3110963		**E. coli* uncharacterized sRNA	HP	–	4,438	1,451	0
sRNA3337545	Both	Bacterial Cobalamin riboswitch	HS	• *cis*-STM3153 [putative inner membrane protein] (5′)*cis*-STM3154 [putative ATP-dependent RNA helicase] (3′)	2,850	1,990	88
sRNA3378366	Plus	–	HP	• *cis-dsbB* [disulfide oxidoreductase; required for periplasmic disulfide bond formation] (3′)	26	0	565
sRNA3405375	Plus	–	HS	• *trans-fklB* [FKBP-type peptidyL-prolyl *cis-trans* isomerase] (5′)*cis-oat* [putrescine aminotransferase] (5′)	0	1,022	1,536
sRNA3551252	Minus	*Sulfolobus acidocaldarius* uncharacterized sRNA	HP	• *cis-yhcO* [Barstar homolog; cytoplasmic inhibitor of extracellular RNAase (Barnase)] (5′)*cis-aaeB* [*p*-hydroxybenzoic acid (*p*-HBA) efflux pump subunit (YhcP)] (3′)	0	199	1,034
sRNA3554557	Both	–	HP	• *cis-aaeX* [membrane protein associated with *p*-HBA efflux pump] (5′)*cis-yhcS* [LysR family transcriptional regulator] (5′)	0	0	201
sRNA3943836	Minus	–	HP	• *cis*-STM3729 [putative Zn^{2+}-binding MPN superfamily protein] (3′)*cis-rpmB* [50S ribosomal protein L28] (5′)	2,718	596	0
sRNA4126333	Plus	*Salmonella enterica* uncharacterized sRNA	HS	• *cis*-rrlC [23S rRNA] (5′)	2,103	436	0
sRNA4130247	Plus	–	HP	• *cis-ilvX* [small (stress response) protein in *E. coli*; not annotated in *S. enterica* genome] (5′)	11,226	5,800	240
sRNA4146790	Both	–	HP	• *cis-trxA* [thioredoxin] (3′)*cis-rho* [transcriptional termination factor Rho] (5′)	2,910	1,268	111
sRNA4184031	Plus	*E. coli* and *Shigella flexneri* sRNA-IS128	HP	• *cis-pldA* [(outer membrane) phospholipase A] (3′)*cis-recQ* [ATP-dependent DNA helicase] (5′)	0	70	184
sRNA4217489	Plus	–	HP	• *cis-hemG* [protoporphyrinogen IX dehydrogenase involved in heme biosynthesis] (3′)*cis-rrsA* [16S rRNA] (5′)	20,763	397	30
sRNA4236850	Minus	**E. coli* sRNA-MicC	HP	• *cis-glnL* [nitrogen regulation protein NR(II); sensor His-kinase in GlnLG 2-component system] (5′)*cis-glnA* [glutamine synthetase] (3′)	71,126	1,674	412
sRNA4250009	Plus	–	HS	–	0	31	1,242

(*Continue on next page*)

Table 3. (*Continued*)

sRNA	Strand[a]	Known sRNA Similarity	Structure[b]	xPutative Gene Targets[Gene Products] / (Overlap)[c]	Exp. In Log[d]	Exp. In 5 -h C-st[d]	Exp. In 24-h C-st[d]
sRNA4385784		–	HS	–	6,191	21,181	100,027
sRNA4415893	Plus	–	HP	• *cis-rrsE* [16S rRNA] (5′)	14,532	2,336	33
sRNA4417787	Plus	–	HS	• *cis-gltV* [tRNA-Glu] (3′)*cis-rrlE* [23S rRNA] (5′)	6085	2,357	40
sRNA4426265	Plus	*E. coli* uncharacterized sRNA	HP	• *cis-aceA* [isocitrate lyase] (3′)*cis-aceK* [bifunctional isocitrate dehydrogenase kinase / phosphatase] (5′)	0	445	0
sRNA4683584	Plus	*E. coli* sRNA-RyhB	HP	• *trans-* STnc2160 [64 bp noncoding RNA] (5′)*cis*-STM4419 (*iolT2*) [MFS superfamily sugar transporter] (3′)	0	59	219
sRNA4720054	Plus	Bacterial tmRNA	HS	• *cis-mgtA* [Mg^{2+} transporting ATPase] (5′)	2,592	1,757	78
sRNA4812322	Plus	*E. coli* uncharacterized sRNA	HS	• *cis-tsr* [methyl-accepting chemotaxis protein I] (3′)*cis*-STM4534 [NtrC family transcriptional regulator] (5′)	3,170	31	0
sRNA4842833	Both	*E. coli* SsrA	HP	• *cis-deoC* [deoxyribose-phosphate aldolase] (3′)*cis-deoA* [thymidine phosphorylase] (5′)	36	196	972

*indicates that there is more than one known sRNA similar to a candidate sRNA in sequence.
[a]Strand – Plus, positive DNA strand;Minus, negative DNA strand; Both, sRNA present on both strands.
[b]Hairpin (HP) – simple sRNA predicted via Mfold; Highly structured (HS) – possesses more than one hairpin.
[c]Gene Target, Gene Product and Overlap were determined from http://www.ncbi.nlm.nih.gov/pubmed Gene and Protein search databases; (5′): indicates that sRNA overlaps 5′-end of CDS; (3′): indicates that sRNA overlaps 3′-end of CDS; (AS): indicates that sRNA is transcribed internally to CDS on the opposite strand; (—) indicates that sRNA is not *cis* (overlapping) to any known gene and does not have an identified *trans* target.
[d]Exp. in Log – transcripts count at sRNA site in non-starved logarithmically growing bacteria; Exp. in 5-h C-st – transcripts count at sRNA site in 5-h C-starved bacteria; Exp. in 24-h C-st – transcripts count at sRNA site in 24-h C-starved cells.

Several candidate sRNAs are related to known sRNAs in other bacteria

As some of our predicted sRNAs would be expected to resemble known bacterial sRNAs, we next aligned our candidates to 3 non-coding RNA databases: Rfam, sRNATarBase and BSRD.[24-27] Rfam is a repository for non-coding RNA families, whereas sRNATarBase and BSRD exclusively house bacterial sRNAs. We found that 43 of the 63 identified sRNAs aligned significantly with known sRNAs and/or RNA elements from other bacterial species (Fig. 3, Supplementary File 1). As indicated in Figure 3 and Supplementary File 1, many of our novel sRNAs share 100% or nearly 100% sequence identity across portions of their sequences to characterized sRNAs or other annotated noncoding RNA elements in other species. That said, it is important to note that we required alignments to meet one of 3 criteria in order to establish relatedness: 90% identity over 35bp, 80% identity over 40bp, or 70% identity over 60bp. While the majority of our alignments meet these criteria, we find none of our full length sRNAs perfectly aligned to any other bacterial genomes whereas each fully aligned to that of SL1344 confirming each of our sRNAs was specifically expressed from SL1344 and not from an experimental contaminant. In addition, structural analysis of the candidate sRNAs using Mfold, suggests >70% of our sRNA candidates assume complex conformations similar to those commonly associated with known bacterial sRNAs (Fig. 4; Supplementary File 2).[10,27]

Furthermore, since many of our novel sRNAs bore significant identity to annotated sRNAs in other species (Fig. 3;

Figure 3. Putative sRNA alignments with other bacterial genomic sequences. (**A**) sRNA523910. (**B**) sRNA1799950. (**C**) sRNA4385784. Alignments were generated using ClustalW. *, 100% nucleotide identity. -, gap.

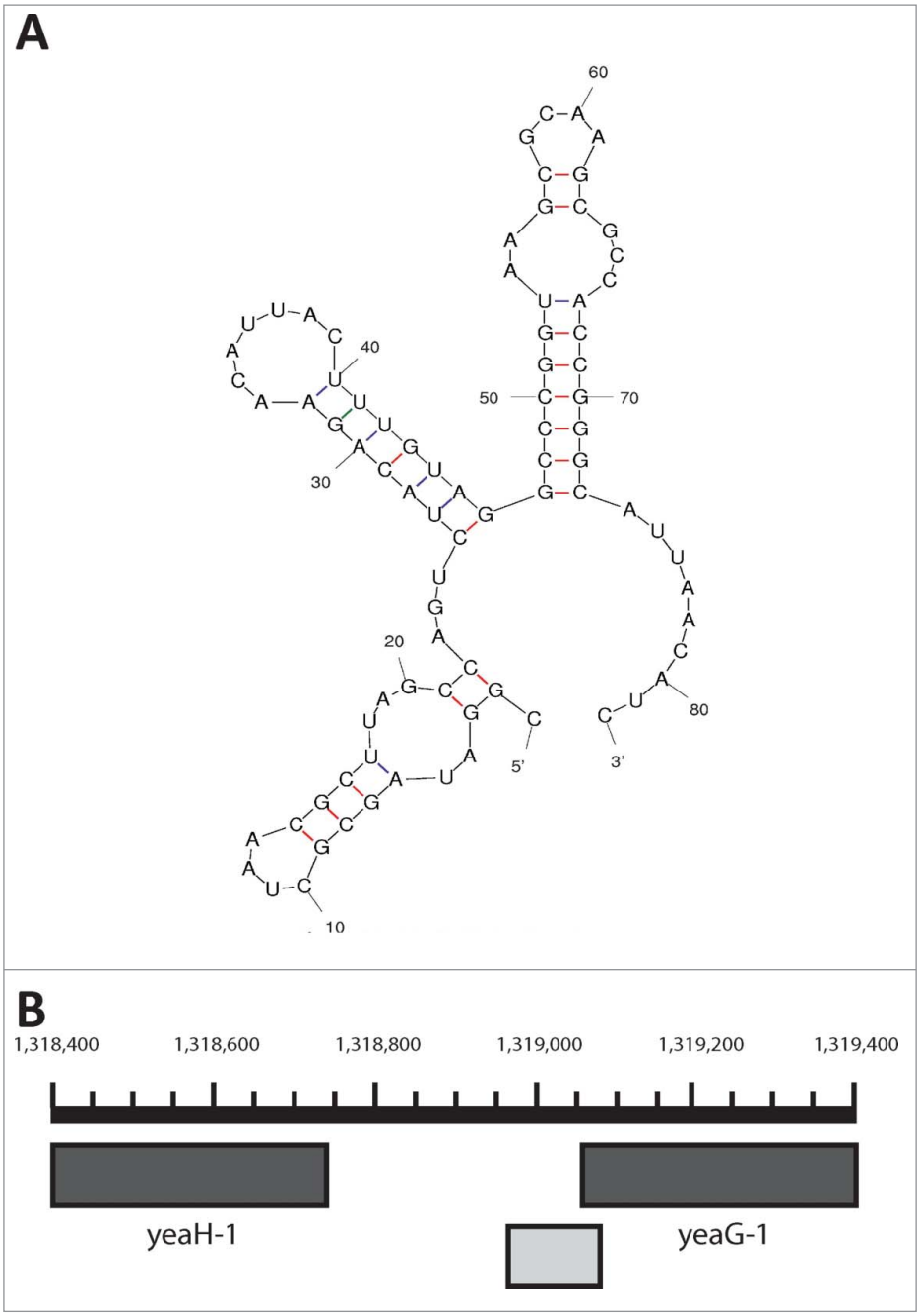

Figure 4. sRNA1318601. (**A**) The Mfold predicted secondary structure of sRNA1318601. (**B**) Diagram depicting the location of sRNA1318601 on genome of *Salmonella enterica* subspecies *enterica* serovar Typhimurium str. SL1344. Dark gray indicates protein coding gene, light gray indicates sRNA.

Supplementary File 1), we decided to examine these related, characterized RNAs for additional similarities and insights into probable functions. Of note, we find homologous genomic positions are generally similarly arranged; typically located intergenically but in close proximity to and often somewhat overlapping with protein coding loci (data not shown). That said, although we find no studies directly examining the responsiveness of any of these homologous sRNA loci to carbon starvation in other species, excitingly, we do find one study[28] that did identify a sRNA (CrfA) in *Caulobacter crescentus* specifically induced upon carbon starvation which does bare ~87% identity to a 34 nt portion of our putative sRNA3551252. While this alignment is below what we generally required for relatedness (Fig. 3 and Supplementary File 1), it is tempting to speculate that these 2 sRNAs may actually be homologous as both are strongly induced upon C starvation and both neighbor related transcriptional activators.

Alignment of candidates with putative targets

As sRNAs function via binding to target mRNAs, we would expect our candidate sRNAs to have sequence complementarity to known SL1344 mRNAs. Strongly supporting the functional relevance of our candidate sRNAs, we find 57 of the 63 putative sRNAs either overlap the 5′ or 3′ ends of neighboring coding DNA sequences (CDS): mRNA, rRNA or tRNA loci; or share

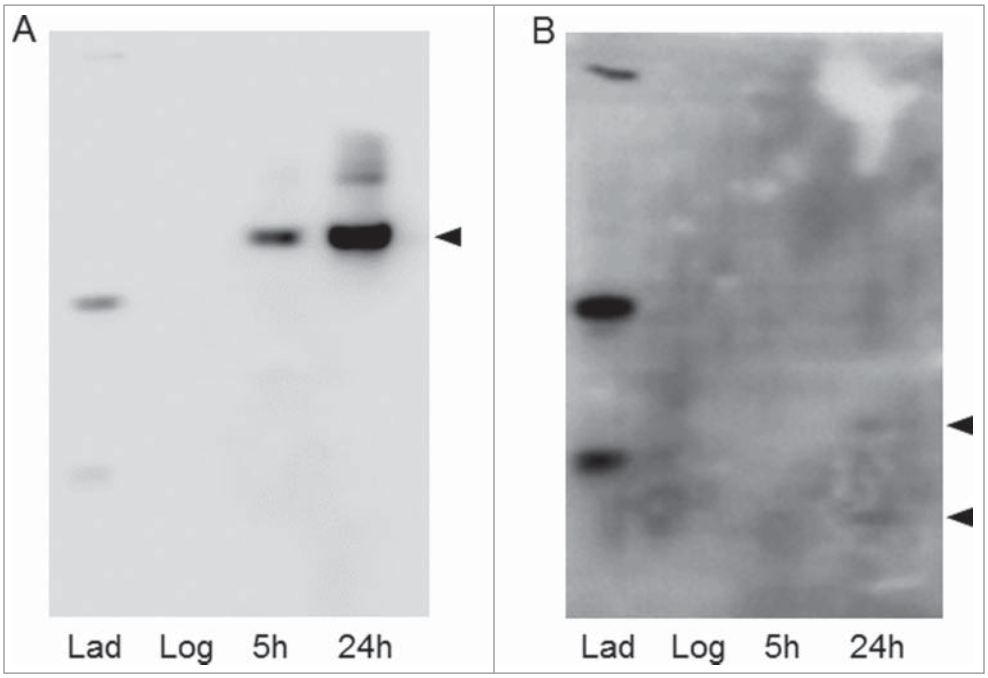

Figure 5. sRNA small transcript RNA gel blots. (**A**) sRNA523910 small transcript northern blot. (**B**) sRNA2838164 small transcript RNA gel blot. Arrowhead indicates predominant sRNA transcript. Lad, custom small DNA 5′ biotinylated ladder with 25, 50 and 200 nt oligonucleotides. Log, logarithmic phase. 5h, 5 hour carbon starved. 24h, 24 hour carbon starved.

significant sequence identities to known mRNAs transcribed elsewhere in the SL1344 genome suggesting the function of the majority of these sRNAs is to regulate transcripts via antisense base-pairing. Of note, we found that all 63 candidate sRNAs are within 300 bp of known SL1344 genes (Supplementary File 2), and while most of the adjacent genes code for mRNA transcripts, some harbor rRNA and / or tRNA genes. In all, our putative target genes include 57 *cis*-encoded and only 4 in *trans*.

Not surprisingly, based upon the growth conditions employed to generate the cell populations from which our RNA was isolated, the majority of the putative targets for the candidate sRNAs identified here are involved in either some aspect of carbon-source metabolism / uptake (approx. 24% of the targeted genes) or stress responses / protection (approx. 27% of targeted genes). The next 2 most targeted groups of genes are components of the translational / protein synthesis machinery including tRNAs, rRNAs and ribosomal proteins (approx. 13% of the targeted genes); as well as known and putative regulators of transcription or translation (approx. 12% of the targeted genes). Most of the remaining putative targeted genes encode gene products involved in synthesis of bacterial structural components (approx. 6.5%) or hypothetical / unknown function proteins (approx. 16%). Interestingly, 6 of the identified sRNAs did not exhibit complementarity to any known mRNA or noncoding RNA in the *S.* Typhimurium SL1344 genome.

Small transcript RNA gel blots

While our RNA-Seq analysis reproducibly generated hundreds and more frequently thousands of unique cDNA reads arising from each of our 63 candidate sRNA loci (average number of distinct reads mapped to each locus = 6,469) providing substantial experimental support for each of our novel sRNA transcripts (Table 3), we also elected to perform small transcript northern blots for 2 of our putative sRNAs in order to: (1) definitively confirm sRNA transcripts are being transcribed from the loci identified in this work by a second experimental strategy and (2) to specifically determine approximate lengths of select sRNA transcripts. Figure 5A shows a small transcript RNA gel blot of sRNA523910 largely agreeing with the relative expressions suggested for this sRNA by our NGS analyses (Table 3). As expected, expression of this sRNA appears to be significantly induced after 5-h of C-starvation followed by a several fold increase in expression by 24-h of C-starvation. Importantly, similar to previous characterizations of *S.* Typhimurium sRNAs,[2] semi-log plot analysis suggests the principle transcript arising from this locus is between 75 and 85 nts long. Interestingly, this locus is immediately adjacent to the 4.5S Signal Recognition Particle locus on the opposing strand known to generate ~120 nt transcripts, as such, it is tempting to speculate that longer 4.5S antisense transcripts might be responsible for the higher molecular weight banding observed or a second RNA component of the signal recognition particle as previously reported[29] (Fig. 5A). Similarly, Figure 5B shows a small transcript northern blot of sRNA2838164 also largely agreeing with the relative expressions suggested for this sRNA by our NGS analyses (Table 3). That said, unlike sRNA523910, we find 2 significant, yet considerably shorter, transcripts arising from the sRNA2838164 locus consisting of approximately 20 and 35 nts.

Discussion

There are still many questions concerning the processes through which *S. enterica* respond to and survive carbon-

energy source starvation. This study focused on identifying previously uncharacterized sRNAs that *S.* Typhimurium strain SL1344 transcribes in 3 different cell populations (growth conditions): (non-starved) log-phase cells, 5-h C-starved, and 24-h C-starved cells. Importantly, we now show that SL1344 does in fact differentially express numerous previously uncharacterized small RNA transcripts in response to the conditions to which these 3 cell populations were exposed. Our RNA-Seq results yielded over 19 million reads that when mapped to the reference SL1344 genome allowed us to identify 63 intergenic regions corresponding to 58 specific, uncharacterized small RNA transcripts and 5 previously characterized sRNAs (Table 3) that are dynamically expressed in response to carbon-energy source availability. As we find 5 of the sRNAs we identified have been previously, independently reported, and 43 of the remaining 58 novel sRNA candidates also significantly align to characterized sRNAs or RNA elements in other bacterial species, we suggest these sequence relationships strongly corroborate our methodology and imply that many of our candidate sRNAs likely do represent previously uncharacterized sRNA molecules with potential regulatory roles in one or more phases of the SSR of *Salmonella enterica*.

The putative target genes for many of our identified sRNAs, not unexpectedly, include genes encoding products involved in carbon-energy source metabolism or uptake, *e.g.*, pyruvate dehydrogenase complex components, bifunctional acetaldehyde-CoA / alcohol dehydrogenase, mannose-specific PTS EnzIID component, and galactose/methyl-galactoside ABC transporter binding protein; as well as gene products involved in other parts of metabolism including a putative lysine-specific permease, glutamine synthetase, a putative glutamate synthase and putrescine aminotransferase (Table 3, Supplementary File 2). Although we do not yet know how these sRNAs specifically regulate these target genes, we do know that induction of pathways and transporter systems for alternative carbon-energy sources and other substrates is part of the overall SSR of *S.* Typhimurium (reviewed in [4]). Another feature of the SSR is the decline in the number of ribosomes in the cell from decreased production and increased degradation of ribosome components [4]; thus the finding that tRNAs, rRNAs and ribosomal proteins are putative targets for several of the identified sRNAs is perhaps not too surprising (Table 3, Supplementary File 2). Even though we do not know how, or if, the sRNAs we predict to target rRNAs might function, we hypothesize that they have a positive effect on the expression / stability of these rRNAs since these sRNAs all exhibited the greatest number of reads in log-phase cells with successively lower levels in 5-h and 24-h C-starved cells, respectively, where they would be less needed. If true, this would provide new and exciting insights into the regulation of the translational and protein synthesis machinery in *Salmonella enterica*.

An important function of the SSR is the generation of cross-resistance to other stresses such as oxidative, acidic pH, thermal, osmotic and antimicrobial peptide-induced stress.[4] So it is not unexpected that the largest group of putatively targeted genes are known, or proposed, to play roles in responses to environmental stresses. Although, it remains experimentally unresolved whether the sRNAs identified here regulate the targeted genes that encode for products such as a small heat-shock IbpA/B-like molecular chaperone, heat-shock Zn^{2+}-dependent protease/chaperone HptX, a putative Mn^{2+}-catalase, a glutathionine S-transferase, peptide-methionine sulfoxide reductase, a periplasmic disulfide oxidoreductase and a FKBP-type peptidyL-prolyl-*cis-trans*-isomerase; these protein's functions could be important in helping the cell renature or degrade denatured proteins and resist or repair oxidative damage resulting from the stress. Interestingly, 2 sRNAs identified here putatively target 2 regulators of biofilm formation in *E. coli* and *Salmonella, bssR* and *csgD*.[30,31] A number of other sRNAs have been found to regulate biofilm formation, in part through CsgD; and, nutrient starvation is an inducer of biofilm formation.[31] Again, this strongly agrees with our hypothesis that many of these sRNAs play regulatory roles within the SSR. In addition, several of the sRNAs identified here appear to target genes encoding putative (as yet uncharacterized) regulatory proteins and genes annotated to encode hypothetical proteins or putative inner and outer membrane proteins of unknown function (Table 3, Supplementary File 2). Thus, this work may also potentially lead to the characterization of unknown and uncharacterized components of the SSR that warrant further investigation.

The sequence similarity between these transcripts and genes known or proposed to be involved in the starvation-stress response suggests a real, functional role for the sRNAs identified here in either modulating the SSR or being involved in other uncharacterized regulatory networks, a possibility that seems even more likely given the conserved hairpin secondary structures of our candidate sRNAs (Fig. 4; Supplementary File 2). Furthermore, significant sequence homology between some of the sRNAs identified here and previously known, but many still uncharacterized, sRNAs of other enterobacteria, and to some extent other enteric pathogens, indicates that these sequences are highly conserved. Further supporting potentially important roles for these sRNAs in controlling the responses of these bacteria to environmental stresses, i.e. carbon-energy source starvation, that are relevant to common niches in the encounter in their life cycles.

As mentioned above, 58 of the 63 sRNAs identified here have not been previously described. Importantly, a previous report examining the early stationary-phase following growth in nutrient-rich media recently brought the total number of identified *S.* Typhimurium sRNAs to 140.[2] To our surprise, we find only 2 of our candidate sRNAs were identified in this work. As for the other 3 previously characterized sRNAs, one was previous identified using a strictly computational analysis of the *E. coli* chromosome,[22] one was identified in a study looking for sRNAs regulated by infection-relevant conditions[21] and one was identified as being associated with the Hfq RNA binding protein.[23]

As we find little overlap between the sRNAs involved in SSR and these other conditions, we now suggest the prevalence of sRNAs and sRNA regulation in prokaryotic systems including *Salmonella enterica* remains largely unexplored and significantly underappreciated. In all, ~300 *Salmonella enterica* sRNAs have been previously reported (detailed in the SalCom database http://bioinf.gen.tcd.ie/cgi-bin/salcom.pl?_HL) meaning the 58 novel characterizations reported here now bring that number to just over 350. That said, however, we find there are 3,035 intergenic regions longer than 50 bps within the SL1344

chromosome, and 2,620 of these regions significantly align to at least one of our unique transcriptome reads.

In conclusion, our results indicate that sRNAs likely have many uncharacterized functions in the survival of *Salmonella enterica* serovars in non-host and host microenvironments that warrant further investigation. Ultimately, the bioinformatic prediction of sRNAs requires experimental evaluation in order to be certain of its validity. In the future, we plan to experimentally confirm additional sRNAs identified in this work by RNA gel blotting then begin characterizing the functional roles for many of our candidate sRNAs, particularly those we believe may function in the regulation of members of the SSR. More broadly, this work, along with several other recent reports identifying a wide range of novel prokaryotic sRNAs, strongly suggest that the prevalence and relevance of bacterial sRNAs is only beginning to be uncovered.

Materials and methods

Bacterial strains and media used

The wild type mouse virulent *Salmonella enterica* serovar Typhimurium strain SL1344 (*hisG46*) was the bacterial strain used in this study.

The minimal media used in this study was a MOPS-buffered salts (MS) containing 22.5 mM potassium phosphate and 10 mM ammonium chloride medium. Log-phase cells were generated using MS medium containing 0.4% (w v^{-1}) glucose (MS hiC). C-starved cells were generated using MS medium with 0.03% (w v^{-1}) glucose (MS loC). Histidine was added at a final concentration of 0.2 mM.

Generation of non-starved log-phase and C-Starved cell populations

SL1344 was grown overnight in MS hiC medium at 37°C with shaking overnight (18–24-h). Overnight cultures were diluted 1:100 into fresh MS hiC and fresh MS loC medium, and incubated with aeration/shaking at 37°C. Growth was monitored spectrophotometrically at 600 nm (optical density at 600 nm; OD_{600}). MS hiC cultures were grown to an OD_{600} of 0.3–0.4 to generate log-phase cells. MS loC cultures were grown until OD_{600} stopped increasing due to exhaustion of glucose and then were C-starved for a total of 5-h and 24-h to generate 5-h C-starved and 24-h C-starved cells, respectively.[6]

Isolation and sequencing of SL1344 RNA

Total RNA from the 3 SL1344 cultures (log-phase, 5-h and 24-h C-starved) was isolated with Trizol® (Life Sciences) per standard manufacture protocol and submitted to Otogenetics (http://www.otogenetics.com/) for commercial RNA-Seq on an Illumina MiSeq genome sequencer using a small RNA protocol. For each culture, Otogenetics provided raw paired-end reads corresponding to ~600 million base pairs. Adapter sequences were removed with cutadapt[32] and the reads were mapped against the SL1344 genome using BWA-MEM.[33]

Identification of intergenic regions of interest

The chromosomal positions of all currently known genes were obtained using an annotated SL1344 genome from Ensembl.[34] As described in Fig. 1, All regions longer than 50 nucleotides that did not contain genes were extracted using an in-house Java program, and these sequences were aligned to the reads from the 3 RNA-Seq experiments using BLAST.[35] A "best hit" parameter was used to ensure individual raw reads only aligned to one location within the intergenic regions, and all the resulting alignments were filtered by length (≥40bp) and percent identity (≥ 98%). The number of reads per unique unannotated region was calculated for each growth condition (log, 5-h and 24-h C-starved), and regions with a large number of alignments in either one or 2 of the cultures but not the other(s) were further characterized as regions of interest.

Identification and analysis of novel small RNAs (sRNAs)

Indexed BAM files representing RNA-Seq reads mapped to the annotated SL1344 reference genome were visualized using the graphical viewer Tablet.[36] Mapped reads within regions of interest were seen as individual lines, and identical or overlapping reads were stacked into rows (Fig. 2). Highly expressed (*i.e.*, stacked) reads that mapped to intergenic regions of interest were extracted as text files, and these were consolidated into a consensus sequence representing one candidate small RNAs. BLAST was then used to align these candidate sRNAs to 3 non-coding RNA databases (Rfam, sRNATarBase and BSRD) in order to examine similarity to known RNAs.[24-27] Mfold was used to predict candidate sRNA secondary structure.[37] To determine possible targets, each candidate sRNA (with 50 bp upstream and downstream flanks) was aligned to all known SL1344 mRNAs as characterized by Ensembl.[24,34] All known genes within 300 bp upstream or downstream of a given candidate sRNA were examined for potential sRNA regulation (Supplementary File 2).

Small transcript northern blots

Total RNA from the 3 SL1344 cultures (log-phase, 5-h and 24-h C-starved) was isolated with Trizol® (Life Sciences) per standard manufacture protocol. A 15% acrylamide/bis-acrylamide (29:1) gel containing 8 M urea (48% (w/v)) and 1X TBE was prerun for 30 min at 100 V in a vertical mini- PROTEAN tank (Bio-Rad). Gels were flushed and loaded with 10 mg of total RNA in 2X TBE/Urea sample buffer (Bio-Rad), then run at 200 V until the bromophenol blue dye front reached the gel bottom. As a size reference, 1 μl of pooled, commercially synthesized biotin 5′ end–labeled DNA oligonucleotides (25, 50 and 200 bp each at 1 μM) was also loaded in 2X TBE/Urea sample buffer. After removal from the electrophoresis plates, gels were gently rinsed with water then washed in 0.5X TBE for 5 min on an orbital shaker. After electrophoresis, RNA was electro-transferred (Mini Trans-Blot Electrophoretic Transfer Cell apparatus, Bio-Rad) to Biodyne B Pre-cut Modified Nylon Membranes 0.45 μm (Thermo Scientific) for 2 h at 20 V in 0.5X TBE. After removal from the transfer stack, membranes were gently washed in 1X TBE for 15 min on an orbital shaker,

then UV cross-linked at 1200 kJ (Stratalinker, Stratagene). Prehybridization was performed in North2South® Hybridization Buffer (Thermo Scientific) at 42° C for 30 min, after which 30 ng (per milliliter of hybridization buffer) of each appropriate biotin 5′ end–labeled oligonucleotide was added directly to the hybridization buffer as probe.

probe2838164 5pBio-AGTATGAAGCTGGTCAATGTT
probe523910 5pBio-GGAAGGAAGCAGCCAAGGCAG

Blots were hybridized overnight with gentle rotation at 42° C. Hybridization buffer was removed the following day, and membranes washed and developed using the Thermo Scientific™ North2South® Chemiluminescent Hybridization and Detection Kit per manufacturer instructions then imaged on a LI-COR C-DiGit Chemiluminescent Western Blot Scanner.

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

Acknowledgments

This work was facilitated by the University of South Alabama (USA) College of Arts & Sciences Department of Biology and the USA College of Allied Health Professions Department of Biomedical Sciences.

Funding

Funding for this work was provided in part by NSF CAREER grant 1350064 (GMB) awarded by Division of Molecular and Cellular Biosciences (with co-funding provided by the NSF EPSCoR program), the South Alabama University Committee on Undergraduate Research (UCUR) (GMB) and the Emma Spector Greenfield Fund for Undergraduate Research, Department of Biomedical Sciences in the College of Allied Health Professions (MPS).

References

1. Vogel J. A rough guide to the non-coding RNA world of Salmonella. Mol Microbiol 2009; 71(1):1-11; PMID:19007416; http://dx.doi.org/10.1111/j.1365-2958.2008.06505.x
2. Kroger C, Dillon SC, Cameron aDS, Papenfort K, Sivasankaran SK, Hokamp K, Chao Y, Sittka A, Hebrard M, Handler K, et al. PNAS Plus: The transcriptional landscape and small RNAs of Salmonella enterica serovar Typhimurium. Proc Natl Acad Sci 2012; 109:E1277-E86; http://dx.doi.org/10.1073/pnas.1201061109
3. Gonzalez-escobedo G, Marshall JM, Gunn JS. NIH Public Access. 2012; 9:9-14
4. Spector MP, Kenyon WJ. Resistance and survival strategies of Salmonella enterica to environmental stresses. Food Research International: Elsevier Ltd 2012:455-81; http://dx.doi.org/10.1016/j.foodres.2011.06.056
5. Reeve C, Bockman A, Matin A. Role of protein degradation in the survival of carbon-starved Escherichia coli and Salmonella typhimurium. J Bacteriol 1984; 157:758-63; PMID:6365890
6. Kenyon WJ, Nicholson KL, Rezuchova B, Homerova D, Garcia-del Portillo F, Finlay BB, Pallen MJ, Kormanec J, Spector MP. ??S-Dependent carbon-starvation induction of pbpG (PBP 7) is required for the starvation-stress response in Salmonella enterica serovar typhimurium. Microbiology 2007; 153:2148-58; PMID:17600059; http://dx.doi.org/10.1099/mic.0.2007/005199-0
7. O'Neal CR, Gabriel WM, Turk AK, Libby SJ, Fang FC, Spector MP. RpoS is necessary for both the positive and negative regulation of starvation survival genes during phosphate, carbon, and nitrogen starvation in Salmonella typhimurium. J Bacteriol 1994; 176:4610-6; PMID:8045891
8. Huang HY, Chang HY, Chou CH, Tseng CP, Ho SY, Yang CD, Ju YW, Huang HD. sRNAMap: Genomic maps for small non-coding RNAs, their regulators and their targets in microbial genomes. Nucleic Acids Res 2009; 37:150-4; http://dx.doi.org/10.1093/nar/gkn852
9. Waters LS, Storz G. Regulatory RNAs in Bacteria. Cell: Elsevier Inc. 2009; 136(4):615-28; http://dx.doi.org/10.1016/j.cell.2009.01.043
10. Li W, Ying X, Lu Q, Chen L. Predicting sRNAs and Their Targets in Bacteria. Genomics Proteomics Bioinformatics 2012; 10(5):276-84; http://dx.doi.org/10.1016/j.gpb.2012.09.004
11. Zamore PD, Haley B. Ribo-gnome: the big world of small RNAs. Science (New York, NY) 2005; 309:1519-24; http://dx.doi.org/10.1126/science.1111444
12. Brantl S. Regulatory mechanisms employed by cis-encoded antisense RNAs. Curr Opin Microbiol 2007; 10(2):102-9; PMID:17387036; http://dx.doi.org/10.1016/j.mib.2007.03.012
13. Fozo EM, Hemm MR, Storz G. Small toxic proteins and the antisense RNAs that repress them. Microbiol Mol Biol Rev 2008; 72:579-89, Table of Contents; PMID:19052321; http://dx.doi.org/10.1128/MMBR.00025-08
14. Gerdes K, Wagner EGH. RNA antitoxins. Curr Opin Microbiol 2007; 10(2):117-24; http://dx.doi.org/10.1016/j.mib.2007.03.003
15. Wagner EGH, Altuvia S, Romby P. Antisense RNAs in bacteria and their genetic elements. Adv Genet 2002; 46:361-98; PMID:11931231; http://dx.doi.org/10.1016/S0065-2660(02)46013-0
16. Storz G, Altuvia S, Wassarman KM. An abundance of RNA regulators. Ann Rev Biochem 2005; 74:199-217; PMID:15952886; http://dx.doi.org/10.1146/annurev.biochem.74.082803.133136
17. Sharma CM, Darfeuille F, Plantinga TH, Vogel J. A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. Genes Dev 2007; 21:2804-17; PMID:17974919; http://dx.doi.org/10.1101/gad.447207
18. Storz G, Opdyke JA, Zhang A. Controlling mRNA stability and translation with small, noncoding RNAs. Curr Opin Microbiol 2004; 7(2):140-4; http://dx.doi.org/10.1016/j.mib.2004.02.015
19. Tsai CH, Liao R, Chou B, Palumbo M, Contreras LM. Genome-wide analyses in bacteria show small-RNA enrichment for long and conserved intergenic regions. J Bacteriol 2015; 197:40-50; PMID:25313390; http://dx.doi.org/10.1128/JB.02359-14
20. Tsai CH, Baranowski C, Livny J, McDonough KA, Wade JT, Contreras LM. Identification of novel sRNAs in mycobacterial species. PloS one 2013; 8:e79411; PMID:24244498; http://dx.doi.org/10.1371/journal.pone.0079411
21. Kroger C, Colgan A, Srikumar S, Handler K, Sivasankaran SK, Hammarlof DL, Canals R, Grissom JE, Conway T, Hokamp K, et al. An infection-relevant transcriptomic compendium for Salmonella enterica Serovar Typhimurium. Cell Host & Microbe 2013; 14:683-95; PMID:24331466; http://dx.doi.org/10.1016/j.chom.2013.11.010
22. Rivas E, Klein RJ, Jones TA, Eddy SR. Computational identification of noncoding RNAs in E. coli by comparative genomics. Curr Biol 2001; 11:1369-73; PMID:11553332; http://dx.doi.org/10.1016/S0960-9822(01)00401-8
23. Sittka A, Sharma CM, Rolle K, Vogel J. Deep sequencing of Salmonella RNA associated with heterologous Hfq proteins in vivo reveals small RNAs as a major target class and identifies RNA processing phenotypes. RNA Biol 2009; 6:266-75; PMID:19333007; http://dx.doi.org/10.4161/rna.6.3.8332
24. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. BMC Bioinformatics 2009; 10:421; PMID:20003500; http://dx.doi.org/10.1186/1471-2105-10-421
25. Cao Y, Wu J, Liu Q, Zhao Y, Ying X, Cha L, Wang L, Li W. sRNATarBase: a comprehensive database of bacterial sRNA targets verified by experiments. RNA (New York, NY) 2010; 16:2051-7; http://dx.doi.org/10.1261/rna.2193110
26. Gardner PP, Daub J, Tate JG, Nawrocki EP, Kolbe DL, Lindgreen S, Wilkinson AC, Finn RD, Griffiths-Jones S, Eddy SR, et al. Rfam:

Updates to the RNA families database. Nucleic Acids Res 2009; 37:136-40; http://dx.doi.org/10.1093/nar/gkn766
27. Li L, Huang D, Cheung MK, Nong W, Huang Q, Kwan HS. BSRD: A repository for bacterial small regulatory RNA. Nucleic Acids Res 2013; 41:233-8; http://dx.doi.org/10.1093/nar/gks1264
28. Landt SG, Lesley JA, Britos L, Shapiro L. CrfA, a small noncoding RNA regulator of adaptation to carbon starvation in Caulobacter crescentus. J Bacteriol 2010; 192:4763-75; PMID:20601471; http://dx.doi.org/10.1128/JB.00343-10
29. Liu L, Ben-Shlomo H, Xu YX, Stern MZ, Goncharov I, Zhang Y, Michaeli S. The trypanosomatid signal recognition particle consists of two RNA molecules, a 7SL RNA homologue and a novel tRNA-like molecule. J Biol Chem 2003; 278:18271-80; PMID:12606550; http://dx.doi.org/10.1074/jbc.M209215200
30. Domka J, Lee J, Wood TK. YliH (BssR) and YceP (BssS) regulate Escherichia coli K-12 biofilm formation by influencing cell signaling. Appl Environ Microbiol 2006; 72:2449-59; PMID:16597943; http://dx.doi.org/10.1128/AEM.72.4.2449-2459.2006
31. Mika F, Hengge R. Small RNAs in the control of RpoS, CsgD, and biofilm architecture of Escherichia coli. RNA Biol 2014; 11:1-14; http://dx.doi.org/10.4161/rna.28867
32. Chen C, Khaleel SS, Huang H, Wu CH. Software for pre-processing Illumina next-generation sequencing short read sequences. Source Code Biol Med 2014; 9:8; PMID:24955109; http://dx.doi.org/10.1186/1751-0473-9-8
33. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009; 25:1754-60; PMID:19451168; http://dx.doi.org/10.1093/bioinformatics/btp324
34. Flicek P, Amode MR, Barrell D, Beal K, Billis K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fitzgerald S, et al. Ensembl 2014. Nucleic Acids Res 2014; 42:749-55; http://dx.doi.org/10.1093/nar/gkt1196
35. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990; 215:403-10; PMID:2231712; http://dx.doi.org/10.1016/S0022-2836(05)80360-2
36. Milne I, Bayer M, Cardle L, Shaw P, Stephen G, Wright F, Marshall D. Tablet-next generation sequence assembly visualization. Bioinformatics 2009; 26(3):401-2; PMID:19965881
37. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res2003; 31:3406-15; PMID:12824337; http://dx.doi.org/10.1093/nar/gkg595