Gene Reports 10 (2018) 167–176

Contents lists available at ScienceDirect

Gene Reports

journal homepage: www.elsevier.com/locate/genrep

Identification and functional characterization of bacterial small non-coding RNAs and their target: A review

Waqas Ahmed[a,b], Mian Abdul Hafeez[c,d,*], Sammina Mahmood[e]

[a] College of Life Sciences, Guangzhou University, Guangzhou, P. R. China
[b] State Key Laboratory of Agricultural Microbiology, College of Veterinary Medicine, Huazhong Agricultural University, Wuhan 430070, China
[c] Department of Parasitology, University of Veterinary and Animal Sciences, Lahore 54000, Pakistan
[d] Department of Pathobiology, Ontario Veterinary College, University of Guelph, Guelph, ON N1G 2W1, Canada
[e] Department of Botany, Government College Women University, Faisalabad

ARTICLE INFO

Keywords:
Bacteria
Non-coding RNA
Target
sRNA prediction
RNA genomics

ABSTRACT

A large number of bacterial small non-coding RNAs (sRNAs) have been identified in the past few years. These are known as heterogeneous group of molecules that act by different mechanisms to modulate a wide range of physiological responses. Many of the identified sRNAs, ranging in size from 40 to 500 nucleotides in length, are highly conserved, located in intergenic regions, and/or in-between open reading frames. Functional studies revealed that sRNAs interact with their target mRNAs by antisense mechanisms, usually around their translation start sites, to modulate gene expression at the posttranscriptional level. Although the number of known sRNAs has dramatically increased in recent years, many challenges in the identification and characterization of sRNAs lie ahead. Here, we review the common strategies used for identification of bacterial sRNAs and their mRNA targets, as well as their functional characterization using experimental approaches which will facilitate our understanding of sRNA biology.

1. Introduction

Bacterial small non-coding RNAs (sRNAs) plays critical role as regulator of gene expression in prokaryotes and eukaryotes, and respond to stressful conditions, environmental changes, and pathogenesis (Storz et al., 2011). Additionally, sRNAs have been assumed to facilitate virulence gene expression and survival of several pathogenic bacteria within their hosts (Papenfort and Vogel, 2010). > 900 sRNAs have been identified and characterized, and the importance of sRNAs in diverse functions urge to discover new sRNAs (Huang et al., 2009a). Typically, bacterial sRNAs are 40–500 nucleotides in length, and do not encompass fully expressed open reading frames (ORFs) (Altuvia, 2007; Ahmed et al., 2016a. These molecules are divided into different classes on the basis of mechanism to control their targets such as *trans*-encoded sRNAs, *cis*-encoded sRNAs, protein binding sRNAs and newly identified CRISPR sRNAs (Waters and Storz, 2009). Of them, *trans*-encoded sRNA are most extensively studied and best characterized sRNAs in bacteria, they have ability to modulate mRNA stability as well as translation through regulating gene expression by imperfect base-pairing.

Computational approaches have been widely used for identification of novel bacterial sRNAs and their experimental validation is becoming an important part of sRNA biology. Although they face challenging task such as they usually do not encompass recurring nucleotide motifs, highly conserved between closely related bacterial species, small in size and expressed in 'pathotype-specific' strains. Different webtools have been established on the basis of RNA structure, comparative genomics and statistical analysis of genomic sequences, and are widely used for sRNA prediction (Backofen and Hess, 2010). The existence of novel sRNA genes revealed promoter and terminator prediction, genome annotations, structure conservation, sequence and base composition statistics. Newly developed computational approaches are still assumed prefect in prediction of sRNA gene due to their effectiveness.

In closely related species several new bacterial sRNAs have been identified. The majority of identified sRNAs interact with dedicated target mRNAs at and around their translation start sites, ultimately affecting their translation and/or stability (Gottesman and Storz, 2011). Trans-encoded sRNAs are located far away from their mRNA-encoding genes in genomic locations. sRNA is assumed to regulate expression of more than one mRNA, and majority of mRNA targets are currently unknown. Numerous sRNA–mRNA interactions based webtool algorithms have been established that propose candidate mRNA targets and predict putative interaction site (Mandin et al., 2007; Tjaden et al., 2006). Recently, number of validated sRNAs has been increased extremely and hundreds of sRNA encoding candidate genes yet needs to

* Corresponding author.
E-mail addresses: mhafeez@uoguelph.ca, abdul.hafeez@uvas.edu.pk (M.A. Hafeez).

https://doi.org/10.1016/j.genrep.2018.01.001
Received 21 October 2017; Received in revised form 18 December 2017; Accepted 4 January 2018
Available online 08 January 2018

discover and validate experimentally. A wide range of computational prediction tool and experimental approaches have been implemented to search for, validate, and functionally characterize the mRNA targets of specific sRNAs. However, there are currently no general *in silico* methods to predict the existence of new sequestrator-like sRNAs.

In the current study, we focused on recent approaches used for identification of novel bacterial sRNAs, prediction of mRNAs target and their functional characterization which will facilitates our understanding of sRNA biology.

2. Mechanism employed by bacterial sRNAs for gene regulation

Bacterial sRNAs are 40–500 nt in length and usually not translated. They modulate a variety of functions such as transcription, translation, DNA maintenance or silencing and mRNA stability. These outstanding regulatory functions are attained through numerous mechanisms, including interactions with DNA, changes in RNA conformation, base pairing with other RNAs and protein binding (Waters and Storz, 2009). Recent advancement in computational prediction and their experimental validation leads to identification of several new sRNAs but their functional characterization in different regulatory networks is still limited (Sorek and Cossart, 2010). The basic mechanisms employed by bacterial sRNAs are described as follows.

(i) *Trans*-encoded sRNA molecules are encoded on the chromosome in *trans* location and possesses short imperfect RNA interactions through RNA duplexes with their target mRNAs, such as degradation of RNase E is encompassed by sRNA-mRNA duplex. sRNA base pairing with target mRNA result in repression of protein levels through mRNA degradation, translational inhibition, or both (Saramago et al., 2014). Furthermore, they are involved in sequestering of ribosome-binding site (RBS) by direct blocking; induce several alterations downstream of RBS, and block translation by base-pairing or *via* direct interaction with mRNAs. Additionally, they engage RNA chaperone protein, Hfq, to enable sRNA-mediated regulation considering limited complementarity between sRNAs and their mRNA targets (Brennan and Link, 2007).

(ii) *Cis*-encoded sRNAs are located complementary to their target and encoded on the strand opposite the gene they regulate. They are often located in the untranslated regions (UTRs) and vary greatly in size. They establish firm RNA duplex formation with corresponding gene which in turn affects translation/ribosome-binding, and reorganize the secondary structures that facilitate termination events or mRNA stability as shown in Fig. 1 (Caldelari et al., 2013).

(iii) Alternatively to the RNA–RNA regulation, sRNA directly interacts with regulatory proteins to influence their activity by mimicking and, thus, proficiently compete with RNA or DNA targets. The CsrA/RsmA family regulators (global carbon storage regulator) are best suited example in this context of regulatory interactions. A common feature in this network present in numerous pathogenic and non-pathogenic bacteria is the transcription of one or more sRNAs such as CsrBC or RsmYZ, depending on a two-component system (TCS) (Babitzke and Romeo, 2007).

3. Recent approaches for identification of novel small non-coding RNAs

Recently, sRNAs have been widely identified and recognized in a large number, assigning a specific function of this diverse class of molecules for important tasks. sRNAs contribution range from gene regulation to chromosome maintenance, RNA modification and editing, translocation of proteins and in stability of mRNAs (Eddy, 2001; Storz, 2002). Most techniques are developed in model organism *E. coli,* and outcome obtained are being transferred to other bacterial pathogens (Mattick, 2001).

3.1. Labeling of total RNA

The first sRNAs was identified with metabolic labeling of *E. coli* total cellular RNA using ^{32}P–orthophosphate and separated by gel fractionation method. Desired bands were then cut out from 1D or 2D polyacrylamide gel electrophoresis and sequence were analyzed by nuclease digestion. sRNA identification using this method assumed correct, considering that RNA is abundant, highly visible and excised from gel as a single RNA species. However, sequence determination by fingerprinting and metabolic labeling method is somewhat cumbersome. Although RNA labeling and sequencing techniques have been improved as compared to first sRNA detected by this method, specifically speaking identification *via* RNA isolation is more feasible in upregulated conditions (Altuvia, 2007). RNA labeling techniques can also be used for highly expressed RNA species. Although it has certain disadvantages but more sensitive for identification of Spot 42, 6S RNA, M1 RNA of RNase P and tmRNA (Moller et al., 2002). Some alternative techniques have been discovered but come with same disadvantages such as in *in vitro* end-labeling of total RNA extracts using T4 RNA ligase, g-[^{32}*P*]ATP and T4 polynucleotide kinase (Watanabe et al., 1997).

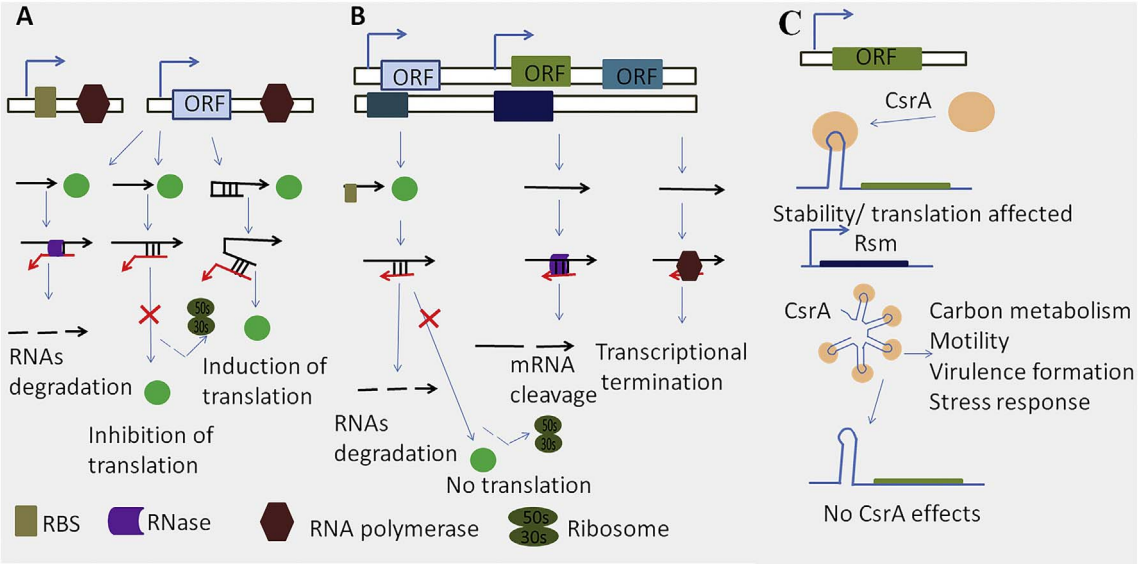

Fig. 1. Simplified representation of mechanisms by which sRNAs function in bacteria. Mechanism of action (A) The *trans*-encoded sRNAs (B) The *cis*-encoded sRNAs (C) RNA-binding protein that interacts with target mRNA (Ahmed et al., 2016a).