

DATABASE

Open Access

sRNAdb: A small non-coding RNA database for gram-positive bacteria

Jordan Pischimarov^{1†}, Carsten Kuenne^{1†}, André Billion¹, Jüergen Hemberger², Franz Cemič², Trinad Chakraborty¹ and Torsten Hain^{1*}

Abstract

Background: The class of small non-coding RNA molecules (sRNA) regulates gene expression by different mechanisms and enables bacteria to mount a physiological response due to adaptation to the environment or infection. Over the last decades the number of sRNAs has been increasing rapidly. Several databases like Rfam or fRNAdb were extended to include sRNAs as a class of its own. Furthermore new specialized databases like sRNAMap (gram-negative bacteria only) and sRNATarBase (target prediction) were established. To the best of the authors' knowledge no database focusing on sRNAs from gram-positive bacteria is publicly available so far.

Description: In order to understand sRNA's functional and phylogenetic relationships we have developed sRNAdb and provide tools for data analysis and visualization. The data compiled in our database is assembled from experiments as well as from bioinformatics analyses. The software enables comparison and visualization of gene loci surrounding the sRNAs of interest. To accomplish this, we use a client-server based approach. Offline versions of the database including analyses and visualization tools can easily be installed locally on the user's computer. This feature facilitates customized local addition of unpublished sRNA candidates and related information such as promoters or terminators using tab-delimited files.

Conclusion: sRNAdb allows a user-friendly and comprehensive comparative analysis of sRNAs from available sequenced gram-positive prokaryotic replicons. Offline versions including analysis and visualization tools facilitate complex user specific bioinformatics analyses.

Background

In recent years numerous small non-coding RNAs (sRNAs) were discovered in bacteria. This class of RNAs is crucial to prokaryotic life, modulating transcription or translation leading to either activation or repression of important physiological processes. sRNAs enable bacteria to trigger rapid physiological responses in order to adapt to the environment or infectious processes [1-3].

To cope with the increasing number of identified sRNAs, databases such as fRNAdb, Rfam, sRNAMap and sRNATarBase were developed [4-9]. All of these approaches have certain drawbacks. fRNAdb contains all classes of RNAs, but allows no further analysis. Rfam is one of the most informative data collections, allowing

detailed analyses via a web front-end. sRNAMap is a webserver-based application for gram-negative bacteria only. sRNATarBase compiles experimental data and allows the prediction of sRNA targets. But all databases available to date limit the analysis to published data only. Therefore bioinformatics analyses of candidate sRNAs in combination with genomes, terminators and other relevant information that has not yet been published is still a very complicated task.

In an attempt to overcome some of the aforementioned drawbacks, we have developed sRNAdb. Our database is a locally installable web-suite, permitting the comparative analysis of sRNAs of gram-positive bacteria including their flanking genes. User modified files in GenBank format and gram-negative bacterial genomes, pooled sRNA candidates or further features of interest can be included in locally installed databases. Furthermore all integrated analysis tools can also be used locally.

* Correspondence: Torsten.Hain@mikrobio.med.uni-giessen.de

[†]Equal contributors

¹Institute of Medical Microbiology, Justus-Liebig-University, Schubertstrasse 81, Giessen D-35392, Germany

Full list of author information is available at the end of the article

Construction and content

A database scheme of unique keys and entities, combined with corresponding relations and connections is given in Figure 1. Optional user defined extensions to locally installed versions of the database are indicated with a lighter background color than the boxes representing database entities.

Input data

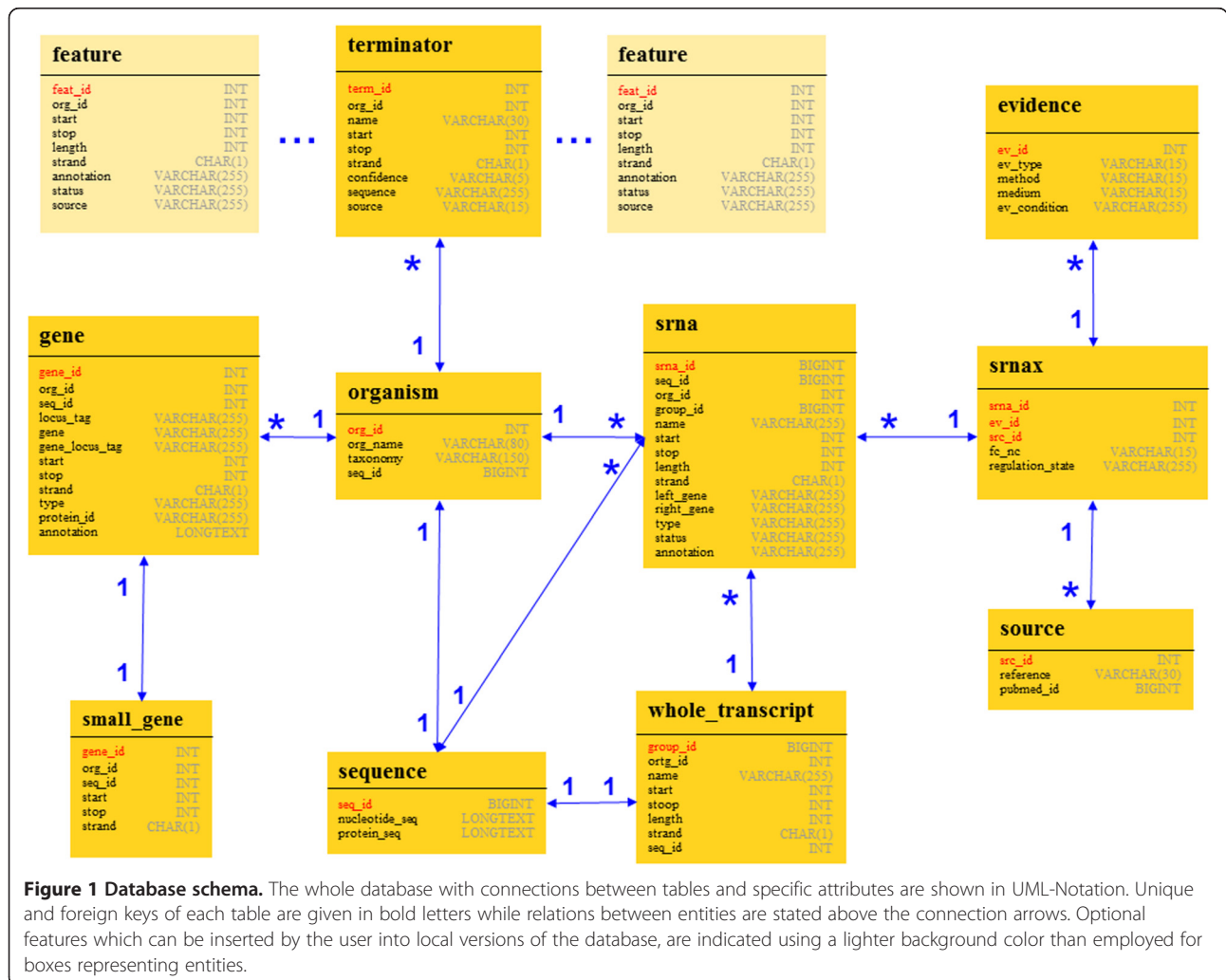
To the best authors' knowledge, no general nomenclature convention for sRNAs exists to date. Therefore sRNAs imported into our database from the literature cannot always be unambiguously distinguished by name, locus or annotation only. Furthermore a large number of published sRNAs is currently annotated as predicted or putative. This leads to a myriad of sRNAs bearing indistinct names, positions or ambiguous annotations. To cope with this difficulty, sRNAdb contains a unique key composed of information about the authors, experimental conditions and sRNA properties as shown in the

table termed srxax of Figure 1. Annotated sequences of organisms or plasmids downloaded from NCBI's RefSeq database [10] represent the replicons in the database. Information annotated in GenBank-formatted files such as sequences, or genes filtered from these files are automatically inserted into sRNAdb. When sRNAdb is installed locally, users can furthermore modify the local database by adding customized features such as terminators, promoters and other additional data. Terminators predicted by TransTermHP [11] serve as examples for this option, as described on the official sRNAdb server homepage.

Architecture and design

Our public sRNAdb server is implemented in Java 1.6 on a Debian Linux PC. It facilitates a client-server architecture using Java Server Pages (JSPs), Java Servlets, and Cascading Stylesheets (CSS). Apache Tomcat and MySQL serve as webserver and database, respectively.

Related sRNAs are determined using BLASTN [12], while protein homologies are established by a combination



of BLASTCLUST and BLASTP [12]. The addition of new data (replicons, sRNAs, terminators, promoters, RBS, etc.) to a local installation of sRNAdb is a simple process based on GenBank and tab-delimited flat-files.

Currently, the public sRNAdb server contains 558 gram-positive genomes and plasmids as well as 9993 automatically predicted and 671 experimentally verified sRNAs. An overview is given in Table 1.

Utility and discussion

The sRNAdb web-database aims to collect all published and predicted sRNAs of gram-positive bacteria for comparative analysis. sRNAs featuring an environmental condition-dependent range of sizes can optionally be joined to a combined transcript. The public version of sRNAdb contains terminators predicted by TransTermHP [11]. Three web-interfaces are provided for retrieval and analysis of the data. The first module is called *search* and offers a rich query interface for the database, as shown in Figure 2A. Properties of sRNAs can be selected and filters can be defined to create task-specific queries resulting in a tabular output (Figure 2B). Related or customized data can also be collated to the query, based on the up- or downstream distance to an sRNA of interest. Furthermore, a secondary structure prediction of selected sRNA sequences by energy minimization can be performed using RNAfold (<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>).

Another interface named *blast* (Figure 3A) was created to enable homology searches of sRNAs versus either

public or proprietary sRNAs or whole chromosomes/plasmids using BLASTN [12]. This can be used for initial screening of potential genomic regions. Concise matrix outputs for comparative analysis purposes as shown in Figure 3B and Figure 3C, are implemented. Complete BLAST alignments are displayed in Figure 3D. Sequences from the BLAST output table can be easily selected by setting checkmarks to extract data into a multifasta-formatted file, ready to serve as input to multiple sequence alignment programs such as CLUSTALW (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>). The resulting output can be used to predict structurally conserved and thermodynamically stable RNA secondary structures using e.g., RNAz (<http://rna.tbi.univie.ac.at/cgi-bin/RNAz.cgi>), facilitating screens for sRNA-homologs across genomes.

For comprehensive visual assessment the *vision* servlet (Figure 4A) was developed. This allows for a comparative analysis of multiple, related chromosome/plasmid loci of the genomic neighborhood of a single sRNA of interest (single mode) as displayed in Figure 4B. The results are translated into an image (.png-formatted) whereby homologous genes (CDS, RNA) of the sRNA locus are identified by BLASTP [12] and presented with an identical colour code. Terminators and any number of additional features previously defined can be included as desired. Each object in the image is associated with a popup-box, displaying further information and linked to corresponding database entries. The width of the resulting image can be varied to compensate for different screen resolutions. Thus one sRNA locus can be

Table 1 The table shows an overview of the current database entries. These are compiled from experiments or from bioinformatic analyses

Reference	sRNAs	Organism	Pubmed_id
Arnvig et al. 2009 [13]	9	<i>Mycobacterium tuberculosis</i> H37Rv	19555452
Bohn et al. 2010 [14]	28	<i>Staphylococcus aureus</i> subsp. <i>aureus</i> N315	20511587
Christiansen et al. 2006 [15]	3	<i>Listeria monocytogenes</i> EGD-e	16682563
D'Hérœul et al. 2011 [16]	22	<i>Enterococcus faecalis</i> V583	21266481
Geissmann et al. 2009 [17]	11	<i>Staphylococcus aureus</i> subsp. <i>aureus</i> N315	19786493
Irnov et al. 2010 [18]	90	<i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168	20525796
Kumar et al. 2010 [19]	50	<i>Streptococcus pneumoniae</i> TIGR4	20525227
Livny et al. 2008 [20]	9993	Gram-positive bacteria	18787707
Mandin et al. 2007 [21]	12	<i>Listeria monocytogenes</i> EGD-e	17259222
Mraheil et al. 2011 [22]	150	<i>Listeria monocytogenes</i> EGD-e	21278422
Nielsen et al. 2008 [23]	1	<i>Listeria monocytogenes</i> EGD-e	18621897
Perez et al. 2009 [24]	33	<i>Streptococcus pyogenes</i> MGAS5005	19888332
Rasmussen et al. 2009 [25]	84	<i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168	19682248
Tezuka et al. 2009 [26]	12	<i>Streptomyces griseus</i> subsp. <i>griseus</i> NBRC 13350	19465662
Toledo-Arana et al. 2009 [27]	103	<i>Listeria monocytogenes</i> EGD-e	19448609
Vockenhuber et al. 2010 [28]	63	<i>Streptomyces coelicolor</i>	21521948

The organisms for which sRNAs are listed in the database, including references, the number of identified sRNAs for the specific organisms and their relevant PubMed identification number are listed.

properties of interest

A

<input checked="" type="checkbox"/> name	<input checked="" type="checkbox"/> start	<input checked="" type="checkbox"/> stop	<input checked="" type="checkbox"/> length
<input checked="" type="checkbox"/> strand	<input checked="" type="checkbox"/> reference	<input checked="" type="checkbox"/> organism	<input type="checkbox"/> sequence
<input type="checkbox"/> status	<input type="checkbox"/> left gene	<input type="checkbox"/> right gene	<input type="checkbox"/> type
<input type="checkbox"/> annotation	<input type="checkbox"/> evidence	<input type="checkbox"/> method	<input type="checkbox"/> medium
<input type="checkbox"/> condition	<input type="checkbox"/> regulation	<input type="checkbox"/> fc/cn	

all/none

set limits

organism

publication

show predicted ncRNAs

↓ set advanced limits

property: <input type="text" value="reference"/>	operator: <input type="text" value="like"/>	value: <input type="text" value="Christiansen"/>	join: <input type="text" value="OR"/>
<input type="text" value="name"/>	<input "="" type="text" value="="/>	<input type="text" value="L13"/>	<input type="text" value="none"/>
<input type="text" value="none"/>	<input type="text" value="none"/>	<input type="text"/>	<input type="text" value="none"/>

show additional features

	upstream [bp] <input type="text" value="20"/>	overlapping <input type="text" value="no"/>	downstream [bp] <input type="text" value="20"/>
terminator	<input type="text"/>	<input type="text" value="no"/>	<input type="text"/>
promoter	<input type="text"/>	<input type="text" value="no"/>	<input type="text"/>
RBS	<input type="text"/>	<input type="text" value="no"/>	<input type="text"/>

search
save
reset
example
?

srna							promoter						
name ↑↓	start ↑↓	stop ↑↓	length ↑↓	strand ↑↓	reference ↑↓	org_name ↑↓	start	stop	length	strand	annotation	status	source
B LhrA	2346166	2346434	269	-	Christiansen et al. 2006	Listeria monocytogenes EGD-e	2346323	2346328	6	-	tataat	predicted	fuzznuc
							2346347	2346359	13	-	tataat_orf	predicted	fuzznuc
							2346354	2346359	6	-	tataat	predicted	fuzznuc
							2346356	2346368	13	-	tataat_orf	predicted	fuzznuc
							2346363	2346368	6	-	tataat	predicted	fuzznuc
							2346381	2346393	13	-	tataat_orf	predicted	fuzznuc
							2346442	2346447	6	-	tataat	predicted	fuzznuc
L13	2685181	2685233	53	-	Toledo-Arana et al. 2009	Listeria monocytogenes EGD-e							
LhrB	1859877	1860017	141	-	Christiansen et al. 2006	Listeria monocytogenes EGD-e	1860017	1860029	13	-	tataat_orf	predicted	fuzznuc
							1860024	1860052	29	-	sig70	predicted	fuzznuc
							1860024	1860029	6	-	tataat	predicted	fuzznuc
							1860029	1860034	6	-	tataat	predicted	fuzznuc
LhrC	232492	232606	115	+	Christiansen et al. 2006	Listeria monocytogenes EGD-e							

Figure 2 Search servlet. Properties of interest for each sRNA such as name, start, stop and so forth can be selected by setting check marks in the *properties* section of the servlet form. sRNAs of specific organisms or publications can be selected according to settings defined in the *set limits* section. Furthermore advanced limits for detailed filtering are available. Additional features like promoters and terminators can be searched for in the neighborhood of sRNAs of interest. **B** An example output from the *search* servlet. The resulting table contains four sRNAs named LhrA, LhrB, LhrC and L13. The corresponding search options are shown in **A**. For each sRNA, properties as well as additional features (promoters) in the surrounding area are displayed in intervals of 20 bp. Also the properties as selected with the *search* servlet are included in the output.

A

query fasta sequence

```
>Listeria_monocytogenes_EGD-e:LhrA
AGGATAGAGTTTACCTCGATGAAAGTTTCATTTTACTAATGTGATTTTCATGGGGATAAGACCCTATCATGGTTATAATGACTTT
TGGTGTATCCATCTTCCAATAGAATGGCACTAGTCAACATCTGTTTAGACACAAAATCATTATTGTTTTCATTTCATCTATTG
GCCGTGTAGTTTCATGTCTATGCCTCATCCCGTTATGCATAGGCGCCGGAACATGATGGCAACGCTGGAAATGTTTATTCGAGCGTC
TTTTTTIG
>Listeria_monocytogenes_EGD-e:L13
GGTATTGTTGGCCCAACAATAAGCCCGGAAGGTTGTTGTTTAAACAATAA
>Listeria_monocytogenes_EGD-e:LhrB
GGTTATTAAACAAGAAGAAGAAGCAACCCGTTCTCGCTCGTTAACACGAACGTTTCAGGCAAAAAATTCAAACTTTCGTGGCG
```

subjects: genomes (dropdown), available target replicons (list), selected target replicons (list)

cutoffs: BLAST (e-value: 10, match/mismatch: 2/-3), BOA (checked), complete matrix (unchecked), identity: 60, query-coverage: 80-180, subject-coverage: 0-100

blast save reset example

B

queries	Listeria_innocua_Clip11262 hits(8)	Listeria_monocytogenes_EGD-e hits(8)	Listeria_seeligeri_serovar_1-2b_str_SLCC3954 hits(7)	Listeria_welshimeri_serovar_6b_str_SLCC5334 hits(8)
Listeria_monocytogenes_EGD-e:L13 (53bp) hits(4)	x	x	x	x
Listeria_monocytogenes_EGD-e:LhrA (208bp) hits(4)	x	x	x	x
Listeria_monocytogenes_EGD-e:LhrB (141bp) hits(4)	x	x	x	x
Listeria_monocytogenes_EGD-e:LhrC (115bp) hits(19)	x(5)	x(5)	x(4)	x(5)

C

query	subject	identity	alignment length	q.coverage	s.coverage	score	e-value
Listeria_monocytogenes_EGD-e:L13	Listeria_innocua_Clip11262	100.0	53	100.0	0.0	96.9	2e-21
Listeria_monocytogenes_EGD-e:L13	Listeria_monocytogenes_EGD-e	100.0	53	100.0	0.0	96.9	2e-21
Listeria_monocytogenes_EGD-e:L13	Listeria_seeligeri_serovar_1-2b_str_SLCC3954	100.0	53	100.0	0.0	96.9	2e-21
Listeria_monocytogenes_EGD-e:L13	Listeria_welshimeri_serovar_6b_str_SLCC5334	100.0	53	100.0	0.0	96.9	2e-21
Listeria_monocytogenes_EGD-e:LhrA	Listeria_innocua_Clip11262	97.77	269	100.0	0.0	455.0	2e-128
Listeria_monocytogenes_EGD-e:LhrA	Listeria_monocytogenes_EGD-e	100.0	269	100.0	0.0	486.0	1e-137
Listeria_monocytogenes_EGD-e:LhrA	Listeria_seeligeri_serovar_1-2b_str_SLCC3954	94.8	269	100.0	0.0	423.0	1e-118
Listeria_monocytogenes_EGD-e:LhrA	Listeria_welshimeri_serovar_6b_str_SLCC5334	97.77	269	100.0	0.0	459.0	2e-129
Listeria_monocytogenes_EGD-e:LhrB	Listeria_innocua_Clip11262	99.29	141	100.0	0.0	250.0	7e-67
Listeria_monocytogenes_EGD-e:LhrB	Listeria_monocytogenes_EGD-e	100.0	141	100.0	0.0	255.0	2e-68

D

query: [Listeria monocytogenes EGD-e:L13](#) and subject: [Listeria innocua Clip11262_start 2776080_end 2776132](#)

```
Query= Listeria_monocytogenes_EGD-e:L13
(53 letters)
>Listeria_innocua_Clip11262
Length = 53

Score = 96.9 bits (106), Expect = 2e-21

Identities = 53/53 (100%)
Strand = Plus / Minus

Query: 1 ggtattgtttgccccacaataagccccggaaggtgtgtgtttaacaataa 53
|||||
Sbjct: 53 ggtattgtttgccccacaataagccccggaaggtgtgtgtttaacaataa 1
```

Figure 3 Blast servlet form and corresponding output. **A** FASTA formatted sRNA sequences can be inserted into the query box. Also target genomes or sRNAs have to be selected for multiple alignment using BLAST. For a detailed BLAST analysis the BLAST output analysis (BOA) options has to be selected. In this example four sRNAs resulting from a search with parameters shown in Figure 1 were selected as input. Genomes of the genus *Listeria* were set as targets and the BOA options were enabled. **B** The number of sRNAs detected in the target organism is displayed in a comparative matrix form. **C** All hits listed in a table and are linked to their corresponding alignment. **D** A detailed BLAST alignment of all results can also be plotted.

A

reference

mode
 sRNAdb
 sequence
 position
 matrix

sRNAdb
 singular whole transcript

Listeria monocytogenes EGD-e
 LhrC-4, LhrC

comparison genomes

available target replicons

Acidimicrobium ferrooxidans DSM 10331
 Acidothermus cellulyticus 11B
 Actinosynnema mirum DSM 43827
 Alicyclobacillus acidocaldarius subsp. acidocalc
 Alicyclobacillus acidocaldarius subsp. acidocalc
 Alicyclobacillus acidocaldarius subsp. acidocalc
 Alicyclobacillus acidocaldarius subsp. acidocalc
 Alkaliphilus metalliredigens QYMF

all + -

selected target replicons

Listeria innocua Clip11262
 Listeria seeligeri serovar 1-2b str. SLCC3954
 Listeria welshimeri serovar 6b str. SLCC5334

properties

BLAST

e-value: 10
 identity: 60
 match/mismatch: 2/-3
 coverage(min-max): 80 -180

image parameters

image width: 756
 range(min-max): 1000 -1000

other sRNAs:
 other hits:
 hit quality:

features

promoter RBS

draw reset example

B

Gene

Identifier: cysK_Imo0223
 Start: 230840
 End: 231766
 Length: 927
 Annotation:
 Product: hypothetical protein
 Note: highly similar to cysteine synthase

sRNA

Identifier: LhrC-4, LhrC
 Start: 232459
 End: 232606
 Length: 148

BLAST

Query: LhrC-4, LhrC
 Start: 255521
 End: 255668
 Length: 148
 Identity: 94.59
 Coverage: 100

Terminator

Identifier: TERM 108
 Start: 232568
 End: 232598
 Length: 31
 Confidence: 100
 Source: TranstermHP

C organism

Listeria innocua 6a Clip11262
 Listeria monocytogenes 08-5578
 Listeria monocytogenes 4b F2365
Listeria monocytogenes EGD-e
 Listeria seeligeri 1-2b SLCC3954
 Listeria welshimeri serovar 6b str SLCC5334

	LhrC	LhrB	LhrA	L13
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				

→ regulatory RNA A → → → E → → → I → → → M → → → Q → → → U → → →

→ overlapping gene B → → → F → → → J → → → N → → → R → → → V → → →

→ flanking genes C → → → G → → → K → → → O → → → S → → → W → → →

 D → H → L → P → T → X →

Figure 4 Vision servlet forms and result of single and batch mode. Different input options are available. After selecting the sRNA of interest, replicons can be selected for visualization. Options for further analyses based on BLAST, as well as properties relating to the image output can be set. **A** An example relating to the LhrC transcript is displayed. **B** Single mode: the resulting image shows a comparative representation of a single sRNA candidate and flanking genes in selected organisms. Moving the mouse pointer over these, the corresponding properties of each object is shown in a separate popup window. **C** Batch mode: sRNAs displayed in Figure 1 are used as input in this example. The output-matrix indicates occurrence of the sRNA candidates in selected organisms and their directional relationships with respect to their surrounding genes.

compared to different chromosomes/plasmids in a concise image output.

For the genome wide analysis of multiple sRNA loci an additional batch mode is available. Results from an application of this batch mode have already been published by Mraheil and collaborators [22]. In order to permit this global analysis an option was implemented that enables export of the data to an Excel sheet. This contains a visualization matrix (Figure 4C) which indicates the occurrence of the sRNA of interest in the target organism together with its directional relationships of the flanking genes.

The software tool presented here is a valuable extension to existing solutions and will assist in the rapid analysis of large volumes of data to understand the distribution and evolution of sRNAs in bacteria. Compared to other databases the comparative batch mode of sRNAdb's *vision* servlet facilitates analyses such as *in silico* screening for phylogenetic markers, or identification of drug targets related to bacterial sRNAs. As exemplified by Mraheil and colleagues [22] a grouping of sRNAs from pathogenic, apathogenic or non-pathogenic bacterial strains based on the *vision* servlet's result matrix, allows the user to identify sRNAs as putative phylogenetic markers. Specifically, sRNAs found exclusively in pathogenic strains can be identified as drug target candidates. Furthermore after download and local installation of sRNAdb, both the database and the dedicated software tools are available to the user. Since proprietary replicons or putative sRNAs can easily be included into locally installed versions of the database, these may be analysed making use of the full power of sRNAdb's software tools, simplifying detailed analyses of unpublished bacterial replicons or sRNA candidates. To the best of the author's knowledge, this functionality is currently not supported by any other publicly available sRNA database.

Conclusion

sRNAdb offers biologists an easy access and analysis to both proprietary and public data and allows the identification of a core set of sRNAs which can be used as putative drug targets in antimicrobial therapeutic approaches as well as specific sRNAs for potential diagnostic markers for the detection of gram-positive bacteria.

Availability and requirements

The database including documentation and tools for analysis are available free of charge at <http://bioinfo.mikro.bio.med.uni-giessen.de/sRNAdb>.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Designed and implemented the database and related software tools: JP, CK, AB, TH. Analyzed data: JP, CK, AB, FC, TH. Wrote the paper: JP, CK, JH, FC, TC, TH. All authors read and approved the final manuscript.

Funding

This work was supported by grants from the German Federal Ministry of Education and Research (BMBF ERA-NET) Pathogenomics Network to the sncRNAomics project (62080061) to T.H. and the German Centre for Infection Research, Justus-Liebig University Giessen.

Author details

¹Institute of Medical Microbiology, Justus-Liebig-University, Schubertstrasse 81, Giessen D-35392, Germany. ²Institute for Biochemical Engineering and Analytics, University of Applied Sciences Giessen-Friedberg, Wiesenstrasse 14, Giessen D-35390, Germany.

Received: 3 April 2012 Accepted: 1 July 2012

Published: 10 August 2012

References

1. Frohlich KS, Vogel J: **Activation of gene expression by small RNA.** *Curr Opin Microbiol* 2009, **12**:674–682.
2. Mraheil MA, Billion A, Kuenne C, Pischmarov J, Kreikemeyer B, Engelmann S, Hartke A, Giard JC, Rupnik M, Vorwerk S, Beier M, Retey J, Hartsch T, Jacob A, Cemic F, Hemberger J, Chakraborty T, Hain T: **Comparative genome-wide analysis of small RNAs of major Gram-positive pathogens: from identification to application.** *Microb Biotechnol* 2010, **3**:658–676.
3. Waters LS, Storz G: **Regulatory RNAs in bacteria.** *Cell* 2009, **136**:615–628.
4. 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* 2010, **16**:2051–2057.
5. Kin T, Yamada K, Terai G, Okida H, Yoshinari Y, Ono Y, Kojima A, Kimura Y, Komori T, Asai K: **fRNAdb: a platform for mining/annotating functional RNA candidates from non-coding RNA sequences.** *Nucleic Acids Res* 2007, **35**:D145–D148.
6. Mituyama T, Yamada K, Hattori E, Okida H, Ono Y, Terai G, Yoshizawa A, Komori T, Asai K: **The Functional RNA Database 3.0: databases to support mining and annotation of functional RNAs.** *Nucleic Acids Res* 2009, **37**:D89–D92.
7. Gardner PP, Daub J, Tate JG, Nawrocki EP, Kolbe DL, Lindgreen S, Wilkinson AC, Finn RD, Griffiths-Jones S, Eddy SR, Bateman A: **Rfam: updates to the RNA families database.** *Nucleic Acids Res* 2009, **37**:D136–D140.
8. Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR: **Rfam: an RNA family database.** *Nucleic Acids Res* 2003, **31**:439–441.
9. 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**:D150–D154.
10. Pruitt KD, Katz KS, Sicotte H, Maglott DR: **Introducing RefSeq and LocusLink: curated human genome resources at the NCBI.** *Trends Genet* 2000, **16**:44–47.
11. Kingsford CL, Ayanbule K, Salzberg SL: **Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake.** *Genome Biol* 2007, **8**:R22.
12. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215**:403–410.
13. Arnvig KB, Young DB: **Identification of small RNAs in *Mycobacterium tuberculosis*.** *Mol Microbiol* 2009, **73**:397–408.
14. Bohn C, Rigoulay C, Chabelskaya S, Sharma CM, Marchais A, Skorski P, Borezee-Durant E, Barbet R, Jacquet E, Jacq A, Gautheret D, Felden B, Vogel J, Bouloc P: **Experimental discovery of small RNAs in *Staphylococcus aureus* reveals a riboregulator of central metabolism.** *Nucleic Acids Res* 2010, **38**:6620–6636.
15. Christiansen JK, Nielsen JS, Ebersbach T, Valentin-Hansen P, Sogaard-Andersen L, Kallipolitis BH: **Identification of small Hfq-binding RNAs in *Listeria monocytogenes*.** *RNA* 2006, **12**:1383–1396.
16. Fouquier DA, Wessner F, Halpern D, Ly-Vu J, Kennedy SP, Serror P, Aurell E, Repoila F: **A simple and efficient method to search for selected primary transcripts: non-coding and antisense RNAs in the human pathogen *Enterococcus faecalis*.** *Nucleic Acids Res* 2011, **39**:e46.

17. Geissmann T, Chevalier C, Cros MJ, Boisset S, Fechter P, Noirod C, Schrenzel J, Francois P, Vandenesch F, Gaspin C, Romby P: **A search for small noncoding RNAs in *Staphylococcus aureus* reveals a conserved sequence motif for regulation.** *Nucleic Acids Res* 2009, **37**:7239–7257.
18. Irnov I, Sharma CM, Vogel J, Winkler WC: **Identification of regulatory RNAs in *Bacillus subtilis*.** *Nucleic Acids Res* 2010, **38**:6637–6651.
19. Kumar R, Shah P, Swiatlo E, Burgess SC, Lawrence ML, Nanduri B: **Identification of novel non-coding small RNAs from *Streptococcus pneumoniae* TIGR4 using high-resolution genome tiling arrays.** *BMC Genomics* 2010, **11**:350.
20. Livny J, Teonadi H, Livny M, Waldor MK: **High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs.** *PLoS One* 2008, **3**:e3197.
21. Mandin P, Repoila F, Vergassola M, Geissmann T, Cossart P: **Identification of new noncoding RNAs in *Listeria monocytogenes* and prediction of mRNA targets.** *Nucleic Acids Res* 2007, **35**:962–974.
22. Mraheil MA, Billion A, Mohamed W, Mukherjee K, Kuenne C, Pischmarov J, Krawitz C, Retey J, Hartsch T, Chakraborty T, Hain T: **The intracellular sRNA transcriptome of *Listeria monocytogenes* during growth in macrophages.** *Nucleic Acids Res* 2011, **39**:4235–4248.
23. Nielsen JS, Olsen AS, Bonde M, Valentin-Hansen P, Kallipolitis BH: **Identification of a sigma B-dependent small noncoding RNA in *Listeria monocytogenes*.** *J Bacteriol* 2008, **190**:6264–6270.
24. Perez N, Trevino J, Liu Z, Ho SC, Babitzke P, Sumbly P: **A genome-wide analysis of small regulatory RNAs in the human pathogen group A *Streptococcus*.** *PLoS One* 2009, **4**:e7668.
25. Rasmussen S, Nielsen HB, Jarmer H: **The transcriptionally active regions in the genome of *Bacillus subtilis*.** *Mol Microbiol* 2009, **73**:1043–1057.
26. Tezuka T, Hara H, Ohnishi Y, Horinouchi S: **Identification and gene disruption of small noncoding RNAs in *Streptomyces griseus*.** *J Bacteriol* 2009, **191**:4896–4904.
27. Toledo-Arana A, Dussurget O, Nikitas G, Sesto N, Guet-Revillet H, Balestrino D, Loh E, Gripenland J, Tiensuu T, Vaitkevicius K, Barthelemy M, Vergassola M, Nahori MA, Soubigou G, Regnault B, Coppee JY, Lecuit M, Johansson J, Cossart P: **The *Listeria* transcriptional landscape from saprophytism to virulence.** *Nature* 2009, **459**:950–956.
28. Vockenhuber MP, Sharma CM, Statt MG, Schmidt D, Xu Z, Dietrich S, Liesegang H, Mathews DH, Suess B: **Deep sequencing-based identification of small non-coding RNAs in *Streptomyces coelicolor*.** *RNA Biol* 2011, **8**:468–477.

doi:10.1186/1471-2164-13-384

Cite this article as: Pischmarov *et al.*: sRNAdb: A small non-coding RNA database for gram-positive bacteria. *BMC Genomics* 2012 **13**:384.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

