A home for RNA families at RNA Biology

Paul P. Gardner & Alex G. Bateman

Published online: 01 Jan 2009.

Submit your article to this journal ⬀

Article views: 55

View related articles ⬀

Citing articles: 5 View citing articles ⬀

Editorial

A home for RNA families at *RNA Biology*

Paul P. Gardner and Alex G. Bateman

Wellcome Trust Sanger Institute; Wellcome Trust Genome Campus; Hinxton, United Kingdom

We are pleased to announce the RNA families track for *RNA Biology*. This track will provide a forum of short publications detailing the structure, function and sequence conservation for RNA families. We envisage two main types of publication in this track, these will consist of analyses of relatively unstudied RNA families or significant review style updates to relatively well studied families. There will be two extra requirements for publication in this track; Firstly the deposition of an alignment and secondary structure in Stockholm format and secondly the generation or update of a corresponding entry in the online encyclopedia Wikipedia. The first requirement will ensure a much needed archive of well curated sequence alignments and secondary structures for all researchers in the community. The second requirement is, as far as we are aware, a first for any scientific publication. The main reason we think this is a good idea is because a Wikipedia entry is usually among the top few hits from a Google search with a molecular biology keyword. Therefore, we would like to ensure that the RNA relevant information in Wikipedia is both reliable and current. We think that this track will provide an important mechanism by which time will be spent by experts to improve the record. In order to ensure this the Wikipedia update will be reviewed alongside the submitted article. Furthermore, the articles will be published both online and in print, they will be open-access and free (at least in the first years of the track). Only articles containing color figures and more than 4 journal pages will incur a fee.

Researchers that have been in the field for a while will note that this track is similar in spirit to the old Protein Sequence Motifs articles published by *Trends in Biochemical Sciences*.[1] This was a very popular track, but it eventually ran out of high impact protein motifs to publish and was discontinued. The RNA field is unlikely to suffer any such problem in the forseeable future given the current rate of discovery of new and important families and classes of RNAs. Some examples of these are the snoRNAs, miRNAs, piRNAs, CRISPRs, guide RNAs and cis-regulatory RNAs such as riboswitches.

These articles will undoubtably provide a valuable resource for projects such as Rfam.[2] Rfam is a small database of RNA alignments and secondary structures that is primarily used for the annotation of nucleotide sequences. These alignments and structures are sourced from literature; Due to a lack of standards for publishing RNA align-ments and structures often the curators resort to manually typing in the sequence and structure from published figures. Obviously this approach is not going to scale well in an era of comparative genomics, deep sequencing of RNAs and RNA gene prediction tools. Therefore it is likely that Rfam will be able to draw alignments and annotation from the RNA families track. Incidentally, Rfam also draws annotations from Wikipedia,[3] therefore the Wikipedia articles written for this journal can be used directly by the database as well as the community. By bringing the three elements of paper, alignment and wikipedia article we beleive that we can allow authors of new families to have a rapid and high impact in the RNA community.

We would like to invite you to contribute articles detailing your novel conserved RNA structural motif to *RNA Biology*. We will also consider significant updates to existing and well known RNA families. The articles should be kept to less than 4 journal pages in length. Guidelines and some notes to get you started on writing a Wikipedia article and formatting a Stockholm alignment can be found at the following URL: http://www.landesbioscience.com/journals/rnabiology/guidelines.

If you would like an assessment as to the suitability of your RNA family feel free to send a presubmission inquiry to the Editor Paul P. Gardner pg5@sanger.ac.uk outlining the motif and its biological relevence.

References

1. McEntyre JR, Gibson TJ. Patterns and clusters within the PSM column in TiBS, 1992–2004. Trends Biochem Sci 2004; 29:627-33.
2. Gardner PP, Daub J, Tate JG, Nawrocki EP, Kolbe DL, Lindgreen S, et al. Rfam: updates to the RNA families database. Nucleic Acids Research 2009; In press.
3. Daub J, Gardner PP, Tate J, Ramsköld D, Manske M, Scott WG, et al. The RNA WikiProject: community annotation of RNA families. RNA 2008; 14:2462-4.

*Correspondence to: Paul P. Gardner; Wellcome Trust Sanger Institute; Wellcome Trust Genome Campus; Hinxton CB10 1SA UK; Email: pg5@sanger.ac.uk

Guidelines For the RNA Families Track at *RNA Biology*

This track will primarily publish articles describing either: (1) substantial updates and reviews of existing RNA families or (2) novel RNA families based on computational and/or experimental results for which little evolutionary analysis has been published. These articles must be accompanied by STOCKHOLM formatted alignments, including a consensus secondary structure or structures and a corresponding Wikipedia article. Publication in the track will require a short manuscript, a high quality Stockholm alignment and at least one Wikipedia article; Each centered around the RNA in question.

Writing the manuscript:

The article itself should adhere to the format: Abstract, Introduction, Materials and Methods, Results, Discussion and Supplementary Material.

• The **Introduction** should give an overview of the family detailing how and when were the original members identified, what is the function (if known) of the family and what was previously known about the taxonomic distribution of the family.

• The **Materials and Methods** section should detail: the sources of seed sequences and structures; the methods and tools used for homology searches, alignments and secondary structure predictions; the sequence databases searched. Correct version numbers and references should be given where relevant.

• The **Results** section should discuss the new homologues (distinguishing paralogues and orthologues) found in the study, discuss the taxonomic distribution and evolutionary conservation of the families sequence and structure. Suggested figures are secondary structure diagrams, alignments (of the smaller families), phylogenetic trees and sequence logos.

• The **Conclusions/Discussion** section should not repeat in detail data given in the Results section. Emphasize the new and important aspects of the study. Relate observations to other relevant studies. On the basis of your findings (and others'), discuss possible implications/conclusions.

• To be eligible for publication the **Supplementary Material** must contain: (1) a link to a Wikipedia article preferably in a User's space. Upon acceptance this can easily be moved into Wikipedia itself together with a reference to the published article; and (2) a plain text Stockholm formatted file containing an alignment and consensus secondary structure for the family.

Formatting your Stockholm alignment:

The quality and consistency of the submitted alignment will be of primary importance for publication in the RNA families track.

The minimal requirements for a Stockholm alignment are:

1. A header line eg. '# STOCKHOLM 1.0'

2. At least two sequences, each preceded with a unique plain-text sequence ID

3. A secondary structure line, preceded with the tag '#=GC SS_cons ', followed by a secondary structure string. Consensus basepairs are represented by matching parentheses, any of '([{<>]})' type parentheses are acceptable. Unpaired sites can be indicated by any of the characters '.,;'. Pseudoknots should be marked-up by matching upper and lower case english alphabet characters. The 5' nucleotide within the knot should be in uppercase and the 3' nucleotide lowercase.

4. A footer line indicating the end of alignment eg. '// '

Ideally the alignment will be marked up with information distinguishing between experimentally validated seed sequences and computationally determined homologs. The sources and relevant identifiers of each sequence can also be documented. These additional pieces of information can be added using '#=GS <seq_id> <tag> <free-text>' tags.

See the Infernal userguide and/or http://en.wikipedia.org/wiki/Stockholm_format for more information. Also, there are a few handy tools for dealing with Stockholm alignments:

• sreformat is useful for converting between different formats.

• Ralee is a handy alignment editor that supports (RNA) Stockholm format.

• Infernal and HMMer both accept Stockholm formats for homology search and alignment.

```
# STOCKHOLM 1.0
#=GF WK http://en.wikipedia.org/wiki/UPSK_RNA
#=GS tym_vir1  CC  EMBL_ID: AF035635.1/619-641  Turnip
yellow mosaic virus    unpublished
#=GS tym_vir2  CC  EMBL_ID: M24804.1/82-104     Turnip
yellow mosaic virus    unpublished
#=GS tym_vir3  CC  EMBL_ID: J04373.1/6212-6234  Turnip
yellow mosaic virus    published    # REF: Virology. 2004 Mar
30;321(1):36-46.
#=GS tym_vir4  CC  EMBL_ID: M24803.1/1-23       Turnip
yellow mosaic virus    unpublished
tym_vir1    UGAGUUCUCGAUCUCUAAAAUCG
tym_vir2    UGAGUUCUCUAUCUCUAAAAUCG
tym_vir3    UAAGUUCUCGAUCUUUAAAAUCG
tym_vir4    UAAGUUCUCGAUCUCUAAAAUCG
#=GC SS_cons .AAA....<<<<aaa....>>>>
//
```

A Short Guide to Creating Your First Wikipedia Article

At least one stub article (essentially an extended abstract) for the paper should be added to either an author's userspace at Wikipedia (preferred route) or added directly to the main Wikipedia space (be sure to add literature references to avoid speedy deletion). This article will be reviewed alongside the manuscript and may require revision before acceptance. Upon acceptance the former articles can easily be exported to the main Wikipedia space. See below for guidelines on how to do this. Existing articles can be updated in accordance with the latest published results.

1. Create an account for yourself at Wikipedia. Go to your favorite Wikipedia article, e.g., http://en.wikipedia.org/wiki/RNA and click on the "Log in / create account" link at the top of the page. Follow the instructions for creating an account for yourself.

2. Once you have logged in you can go to your userpage by clicking on your username at the top of the page (mine takes me to the URL http://en.wikipedia.org/wiki/User:Ppgardne). You can personalize this page as you like. To create your stub article, simply add "/name_of_RNA_family" to the URL. For example, if I wanted to create an article about the new eRNA family I would visit the URL "http://en.wikipedia.org/wiki/User:Ppgardne/eRNA". Click on the option "Start the User:Ppgardne/eRNA page" and then add the content that you would like to appear in the article to the text box.

You can view previews of the article as you work. Do not forget to click save once you are happy with the article.

• An excellent option for inexperienced users is to write a draft of the article using OpenOffice.org (OO) version 2.3 or greater. OO can export content directly in WikiMedia format!

• Keep in mind that Wikipedia articles are to be targeted at a level that an undergraduate could comprehend. Try to avoid jargon and do provide links to other Wikipedia articles at the first use of specific terms, e.g., [[RNA]]. Also the title of the page should appear in bold at the first use in the text of the article, e.g., '"eRNA"'.

3. References can be generated in Wikipedia format using the template filler tool:

http://diberri.dyndns.org/cgi-bin/templatefiller/index. cgi?ddb=&type=pubmed_id&id=&add_ref_tag=1

Add these at end of relevant sentences within the text after punctuation points.

4. To get a nicely formatted reference list at the bottom of the page just add the following two lines:

==References==
{{reflist|2}}

• For more help see: http://en.wikipedia.org/wiki/Help:Starting_a_new_page

• Take note to ensure that any uploaded figures conform to Wikipedia's image use policy (http://en.wikipedia.org/wiki/Wikipedia:Image_use_policy).

• An example of a good RNA page is the hammerhead ribozyme article (http://en.wikipedia.org/wiki/Hammerhead_ribozyme).