

## **Supplementary Information**

### **RNA Secondary Structure Prediction By Centroids in a Boltzmann Weighted Ensemble**

Ye Ding<sup>¶, #</sup>, Chi Yu Chan<sup>#</sup>, and Charles E. Lawrence<sup>¶, #, \*</sup>

<sup>¶</sup>Correspondence:

Email: yding@wadsworth.org; Phone: (518) 486 1719; Fax: (518) 402 4623

Email: lawrence@dam.brown.edu; Phone: (401) 863 1479; Fax: (401) 863 1355

# Bioinformatics Center, Wadsworth Center, New York State Department of Health  
150 New Scotland Avenue, Albany, NY 12208

\* Center for Computational Molecular Biology, and Division of Applied Mathematics, Brown  
University, 182 George Street, Providence, RI 02912

## Supplementary Tables

**Table S1.** The small subunit (16S or 16S-like) rRNAs used in this study (Cannone *et al.*, 2002).

| <b>Organism</b>                  | <b>GenBank accession number</b> |
|----------------------------------|---------------------------------|
| <i>Aquifex aeolicus</i>          | AE000709 AE000751               |
| <i>Babesia bigemina</i>          | X59604                          |
| <i>Bordetella bronchiseptica</i> | U04948                          |
| <i>Borrelia burgdorferi</i>      | M88329                          |
| <i>Deinococcus radiodurans</i>   | M21413                          |
| <i>Escherichia coli</i> K12      | AE000460                        |
| <i>Haloferax volcanii</i>        | K00421                          |
| <i>Mycoplasma hyopneumoniae</i>  | Y00149                          |
| <i>Pirellula marina</i>          | X62912                          |
| <i>Streptomyces eurythermus</i>  | D63870                          |

**Table S2.** The large subunit (23S or 23S-like) rRNAs used in this study (Cannone *et al.*, 2002).

| <b>Organism</b>                    | <b>GenBank accession number</b>    |
|------------------------------------|------------------------------------|
| <i>Caenorhabditis elegans</i>      | X54252                             |
| <i>Chlamydomonas reinhardtii</i>   | X15727 X16686 X16687               |
| <i>Deinococcus radiodurans</i>     | AE002087                           |
| <i>Deinococcus radiodurans</i>     | AE001886                           |
| <i>Escherichia coli</i>            | J01695                             |
| <i>Haloarcula marismortui rrnB</i> | AF034620                           |
| <i>Neisseria meningitidis</i>      | X67300                             |
| <i>Oryza sativa</i>                | M11585 M16845 M35384 M82048 M19228 |
| <i>Plasmodium falciparum</i>       | X61660 X95275                      |
| <i>Thermus thermophilus</i>        | X12612                             |

**Table S3.** The 5S rRNAs used in this study (Cannone *et al.*, 2002).

| <b>Organism</b>                       | <b>GenBank accession number</b> |
|---------------------------------------|---------------------------------|
| <i>Agrobacterium tumefaciens</i>      | X02627                          |
| <i>Deinococcus radiodurans</i>        | AE002087                        |
| <i>Escherichia coli</i>               | V00336                          |
| <i>Geobacillus stearothermophilus</i> | AJ251080                        |
| <i>Geobacillus stearothermophilus</i> | M25591                          |
| <i>Geobacillus stearothermophilus</i> | M24839                          |
| <i>Haloarcula marismortui</i>         | AF034620                        |
| <i>Rhodobacter capsulatus</i>         | X04585                          |
| <i>Saccharomyces cerevisiae</i>       | X67579                          |
| <i>Thermus thermophilus</i>           | V01415                          |

**Table S4.** The Group I introns used in this study (Cannone *et al.*, 2002). The intron subgroup is indicated for each of the sequences.

| <b>Organism</b>                                      | <b>Subgroup</b> | <b>GenBank accession number</b> |
|--|-----------------|---------------------------------|
| <i>Acanthamoeba griffini</i>                         | IC1             | S81337                          |
| <i>Acanthamoeba griffini</i>                         | IC1             | U02540 U07412                   |
| <i>Bangia atropurpurea</i>                           | IC1             | D88387                          |
| <i>Bangia fuscopurpurea</i>                          | IC1             | AF342745                        |
| <i>Hildenbrandia rubra</i>                           | IC1             | L19345                          |
| <i>Metarhizium anisopliae</i> var. <i>anisopliae</i> | IC1             | AF197120                        |
| <i>Metarhizium anisopliae</i> var. <i>anisopliae</i> | IC1             | AF197122                        |
| <i>Porphyra leucosticta</i>                          | IC1             | AF342746                        |
| <i>Tetrahymena thermophila</i>                       | IC1             | V01416 J01235                   |

**Table S5.** The Group II introns used in this study (Cannone *et al.*, 2002). The intron subgroup is indicated for each of the sequences.

| <b>Organism</b>                 | <b>Subgroup</b> | <b>GenBank accession number</b> |
|---------------------------------|-----------------|---------------------------------|
| <i>Saccharomyces cerevisiae</i> | IIB             | V00694                          |
| <i>Saccharomyces cerevisiae</i> | IIA             | AJ011856                        |

**Table S6.** The RNase P RNAs used in this study (Brown, 1999).

| <b>Organism</b>                    | <b>GenBank accession number</b> |
|------------------------------------|---------------------------------|
| <i>Bacillus halodurans</i>         | AP001513                        |
| <i>Dermocarpa</i> sp.              | X97396                          |
| <i>Desulfovibrio desulfuricans</i> | M59357                          |
| <i>Heliobacterium chlorum</i>      | U64881                          |
| <i>Leptospirillum ferrooxidans</i> | AF296042                        |
| <i>Mycoplasma capricolum</i>       | D13066                          |
| <i>Tarsius syrichta</i>            | L08801                          |
| <i>Uta stansburiana</i>            | AF044325                        |
| <i>Verrocormicribium spinosum</i>  | AF296043                        |
| <i>Zygosaccharomyces bailii</i>    | AF186231                        |

**Table S7.** The signal recognition particle (SRP) RNAs used in this study (Alm Rosenblad *et al.*, 2003; Larsen & Zwieb, 1991).

| <b>Organism</b>   | <b>GenBank accession number</b> |
|---|---------------------------------|
| <i>Arabidopsis thaliana</i> (Variety B)                 | AC013453                        |
| <i>Arabidopsis thaliana</i> (Variety F)                 | X55111                          |
| <i>Chlorobium tepidum</i>                               | AE012781                        |
| <i>Humulus lupulus</i> (Variety A)                      | X65984                          |
| <i>Lycopersicon esculentum</i> (Variety C)              | Z29099                          |
| <i>Methanobacterium thermoautotrophicum</i> (Variety A) | X15364                          |
| <i>Methanopyrus kandleri</i>                            | AE010387                        |
| <i>Methanothermus fervidus</i>                          | M32222                          |
| <i>Micrococcus luteus</i>                               | M31831                          |
| <i>Pyrococcus horikoshii</i>                            | AB009466                        |

**Table S8.** The tRNAs used in this study (Sprinzl *et al.*, 1998). The identification code used in the online database is included for each of the sequences.

| <b>Organism</b>                  | <b>Database ID</b> | <b>GenBank accession number</b> |
|----------------------------------|--------------------|---------------------------------|
| <i>Crossostoma lacustre</i>      | DF5210             | M91245                          |
| <i>Erinaceus europaeus</i>       | DL5552             | X88898                          |
| <i>Lactobacillus delbrueckii</i> | DS1520             | X13888                          |
| <i>Methanothermus fervidus</i>   | DT0680             | M26977                          |
| <i>Oryza sativa</i>              | DL2700             | AE017082                        |
| <i>Pichia guilliermondii</i>     | DL7770             | D17533                          |
| <i>Stigmatella aurantiaca</i>    | DT1630             | X82820                          |
| <i>Synechocystis</i> sp.         | DR2141             | U37695                          |
| <i>Tetrahymena pyriformis</i>    | DN7530             | X16643                          |
| <i>Trypanosoma brucei</i>        | DA3681             | M94286                          |

**Table S9.** The tmRNAs used in this study (Zwieb *et al.*, 2003).

| <b>Organism</b>  | <b>GenBank accession number</b> |
|--|---------------------------------|
| <i>Bacteroides fragilis</i>                              | GSP <sup>a</sup>                |
| <i>Buchnera aphicola</i> strain Sg (Schizaphis graminum) | AE014094                        |
| <i>Caulobacter crescentus</i>                            | AF255738                        |
| <i>Clostridium acetobutylicum</i> (partial sequence)     | AE007587                        |
| <i>Dehalococcoides ethenogenes</i> strain 195            | GSP <sup>b</sup>                |
| <i>Klebsiella pneumoniae</i>                             | GSP <sup>c</sup>                |
| <i>Mesostigma viride</i> chloroplast                     | AF166114                        |
| <i>Mycobacterium leprae</i>                              | Z98271                          |
| <i>Rickettsia prowazekii</i>                             | AJ235273                        |
| <i>Trichodesmium erythraeum</i>                          | GSP <sup>d</sup>                |

<sup>a</sup> [http://www.sanger.ac.uk/Projects/B\\_fragilis/](http://www.sanger.ac.uk/Projects/B_fragilis/)

<sup>b</sup> [http://tigrblast.tigr.org/ufmg/index.cgi?database=d\\_ethenogenes|seq](http://tigrblast.tigr.org/ufmg/index.cgi?database=d_ethenogenes|seq)

<sup>c</sup> <http://genome.wustl.edu/projects/bacterial/kpneumoniae/>

<sup>d</sup> [http://genome.jgi-psf.org/draft\\_microbes/trier/trier.home.html](http://genome.jgi-psf.org/draft_microbes/trier/trier.home.html)