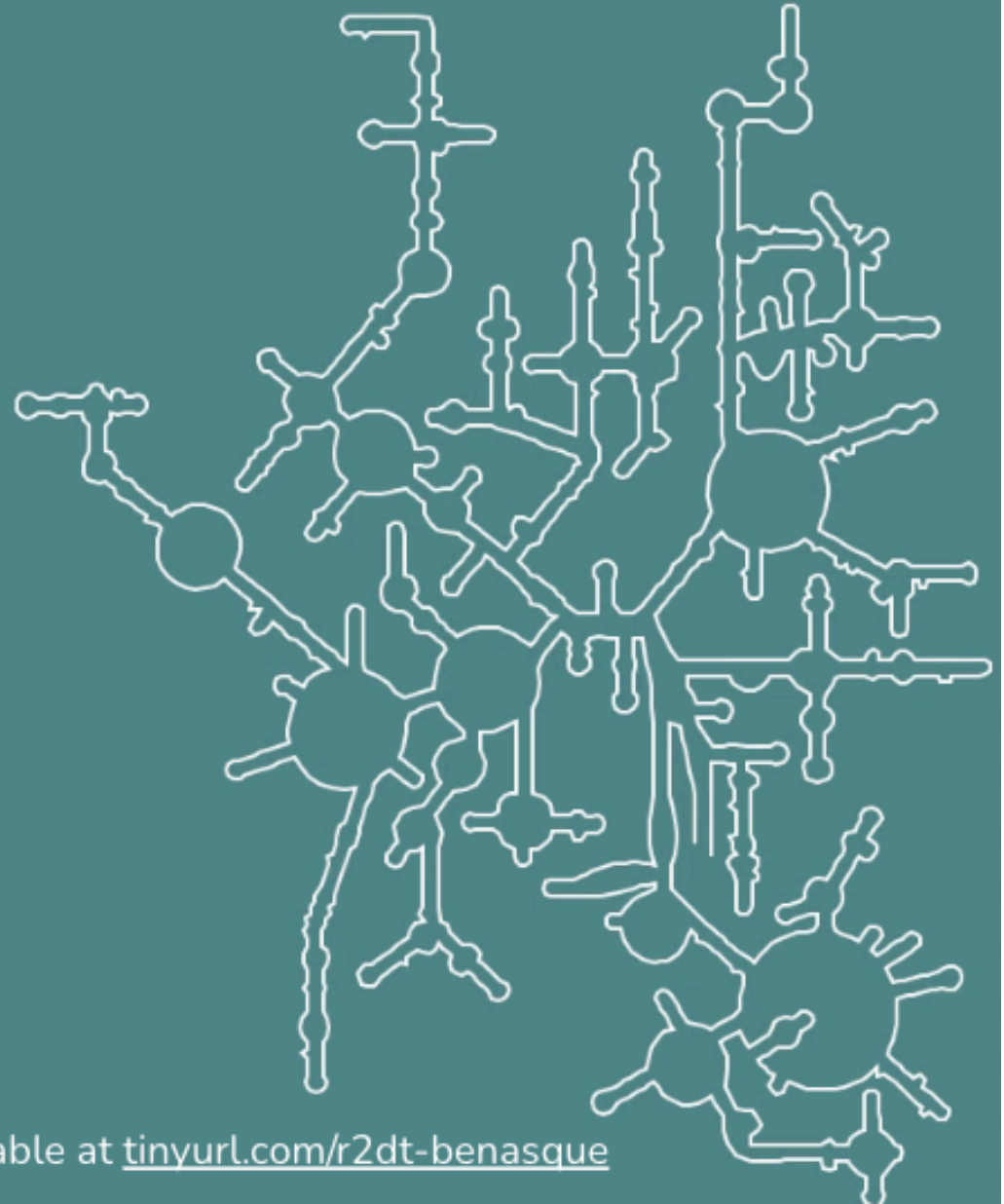


Visualising RNA secondary structure using R2DT

Anton I. Petrov
on behalf of the R2DT team

Aug 19th, 2022 | Benasque

These slides are available at tinyurl.com/r2dt-benasque



The R2DT team

UCSC

Patricia Chan
Todd Lowe

BGSU

Craig Zirbel

NCBI

Eric Nawrocki

Independent

Anton I. Petrov

Georgia Tech

Anton S. Petrov

Holly McCann

Caeden Meade

Aparna Maddala

Loren D. Williams

University of Texas at Austin

Robin Gutell

Jamie Cannone

EMBL-EBI

Carlos Ribas

Fabio Madeira

Mandar Deshpande

Sri Devan Appasamy

Stephen Anyango

Mihaly Varady

Rob Finn

Blake Sweeney

Charles University

David Hoksza

My slide from 4 years ago

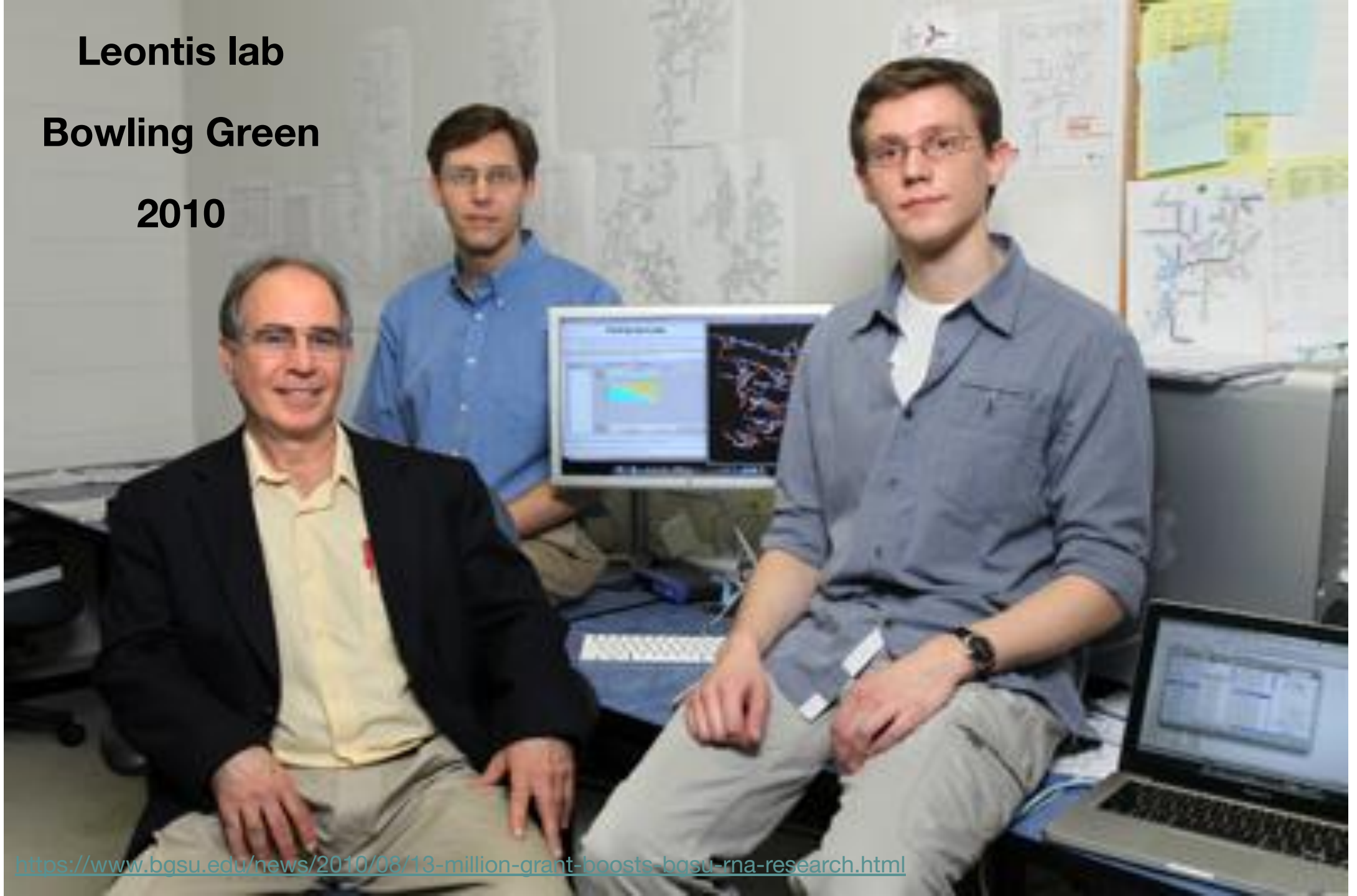
Benasque 2018
top-secret project:

display 2Ds for **all rRNAs**
in RNACentral
using **standard layouts**

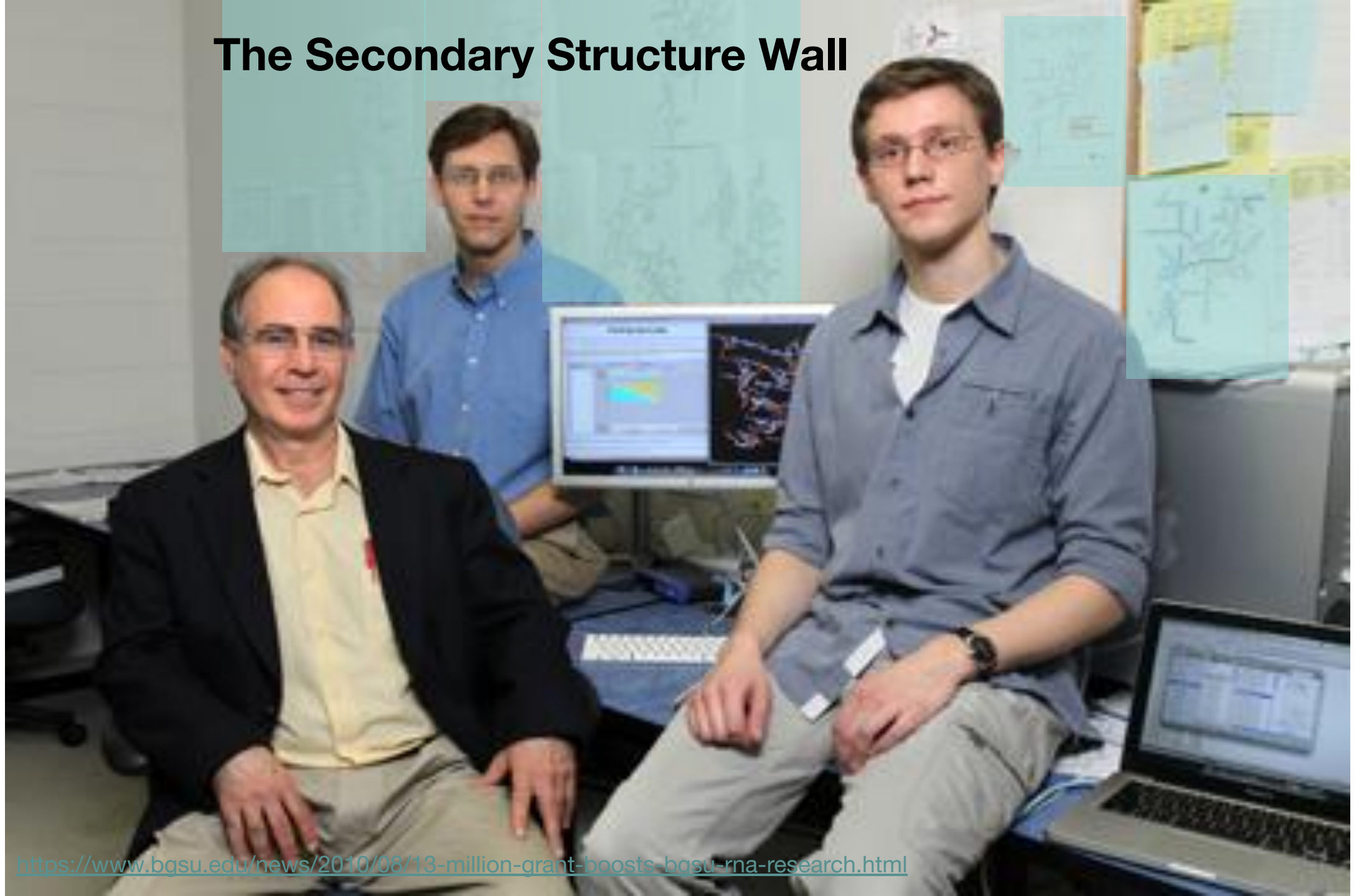
Leontis lab

Bowling Green

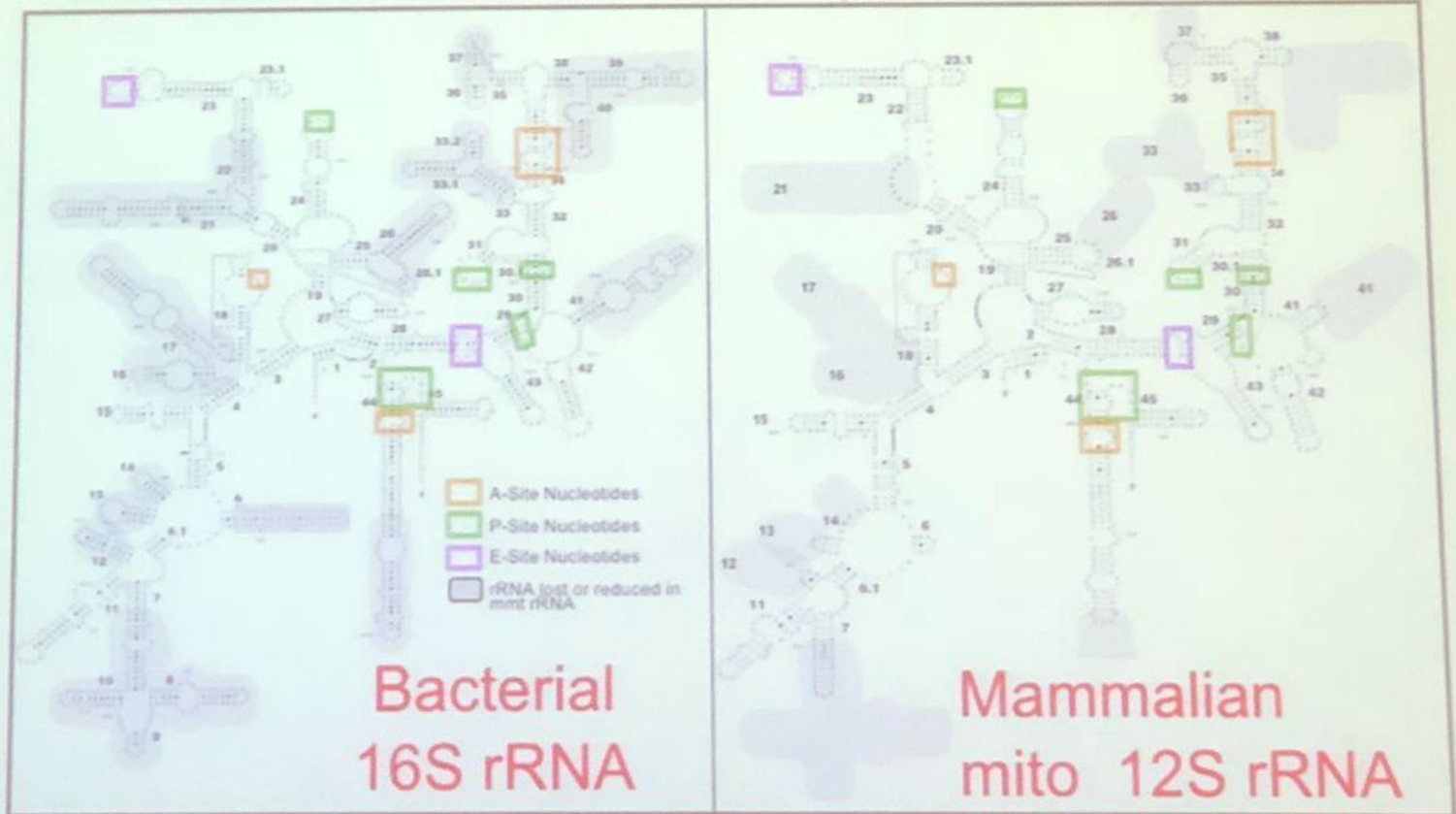
2010

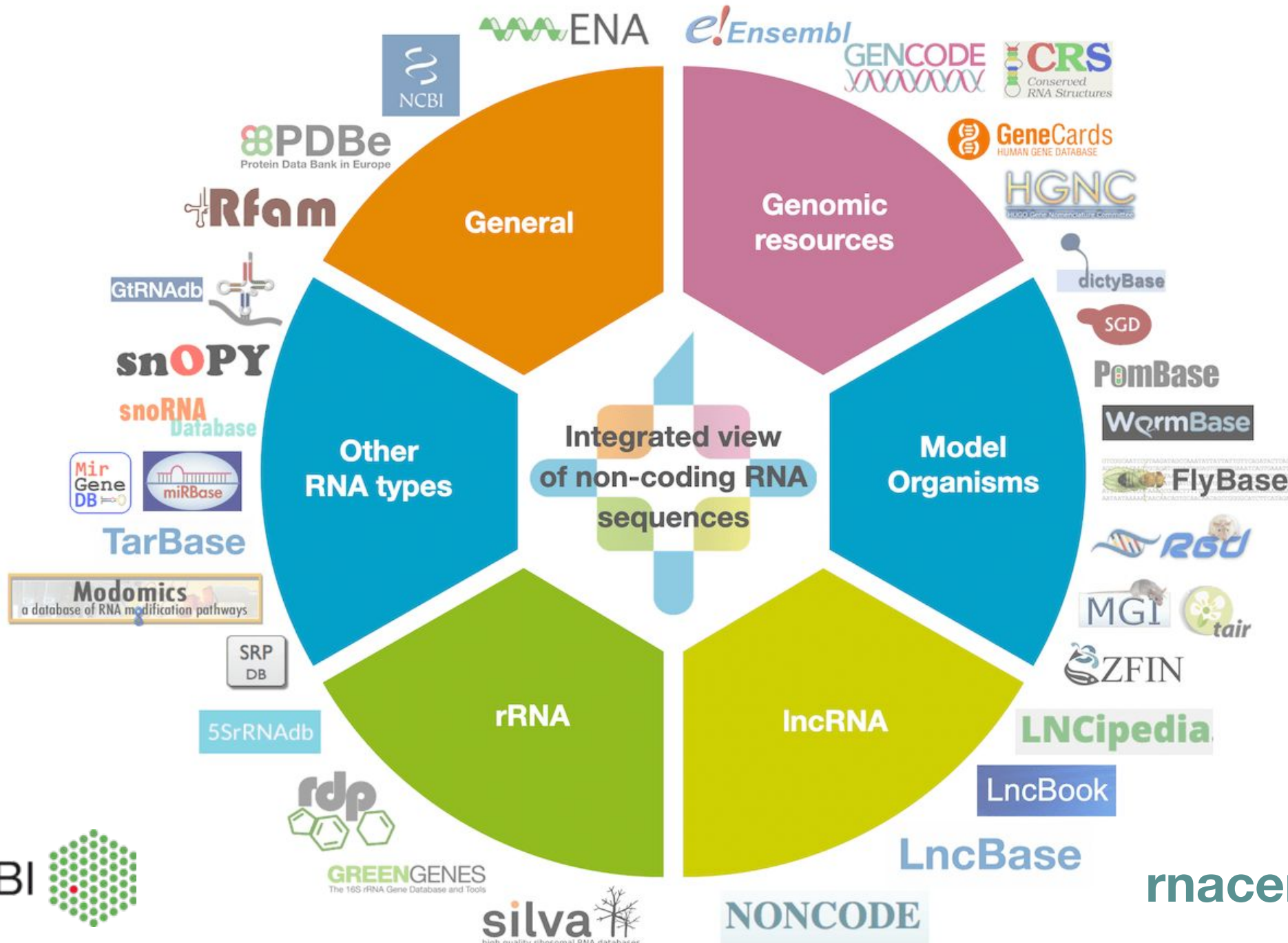


The Secondary Structure Wall

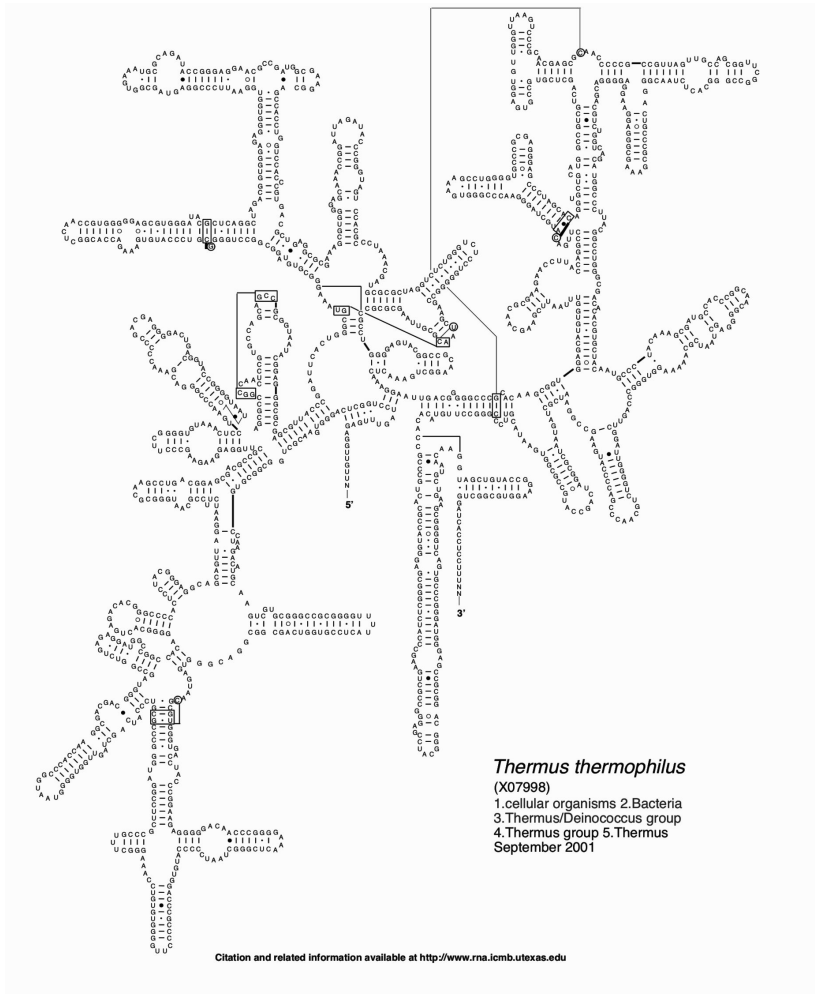


Mmt SSU 12S rRNA: Loss of Peripheral RNA Elements

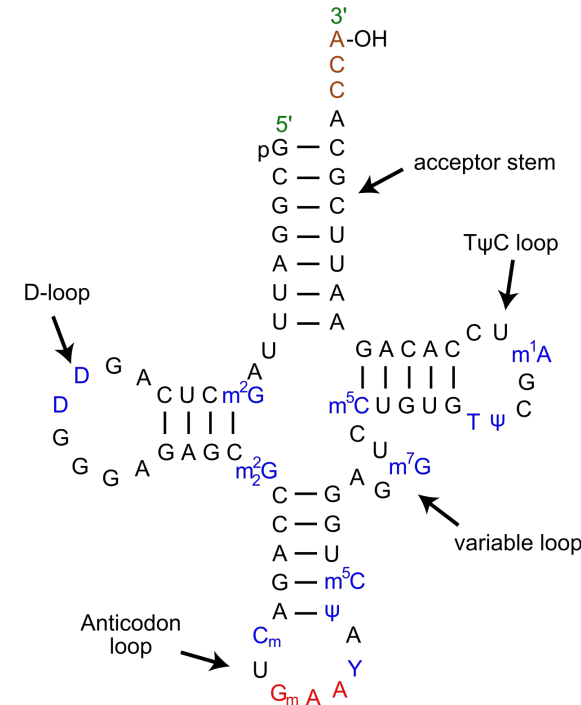




Some RNAs need to be visualised in standard orientations



Ribosomal RNA (rRNA)

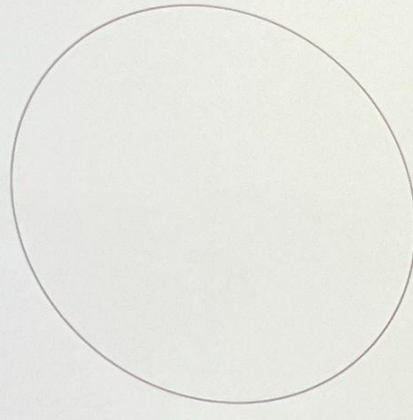


Transfer RNA (tRNA)

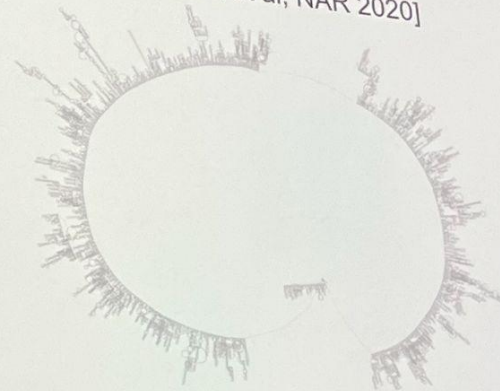
The majority of
RNA central
sequences are
rRNAs and tRNAs

Automated layouts lack in grace

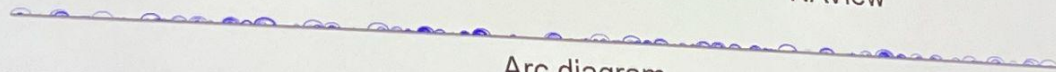
In vivo SHAPE model (bounded BP span) [Manfredonia et al, NAR 2020]



Circle layout



NAView

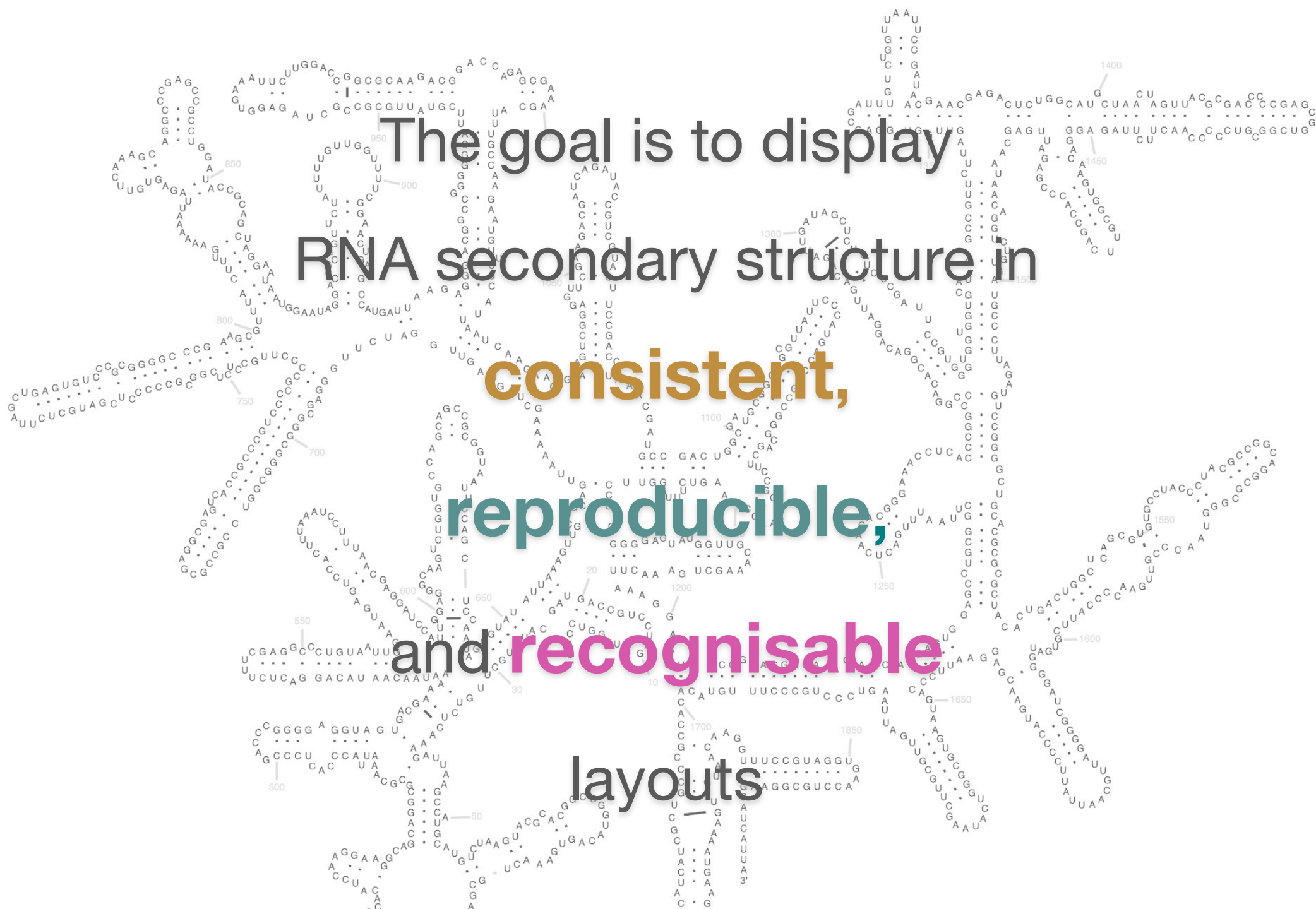


Arc diagram



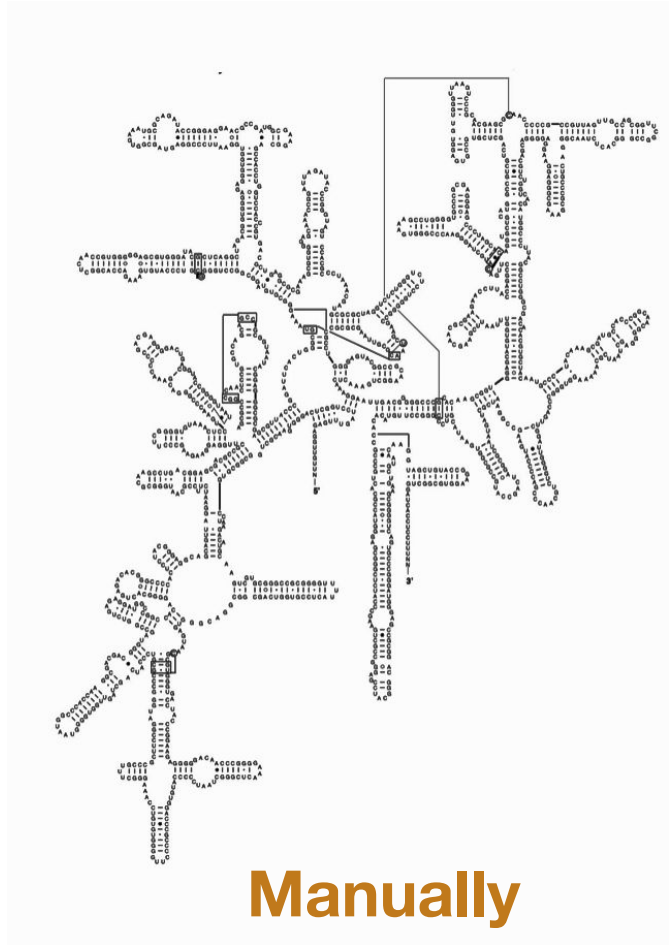
Radial layout



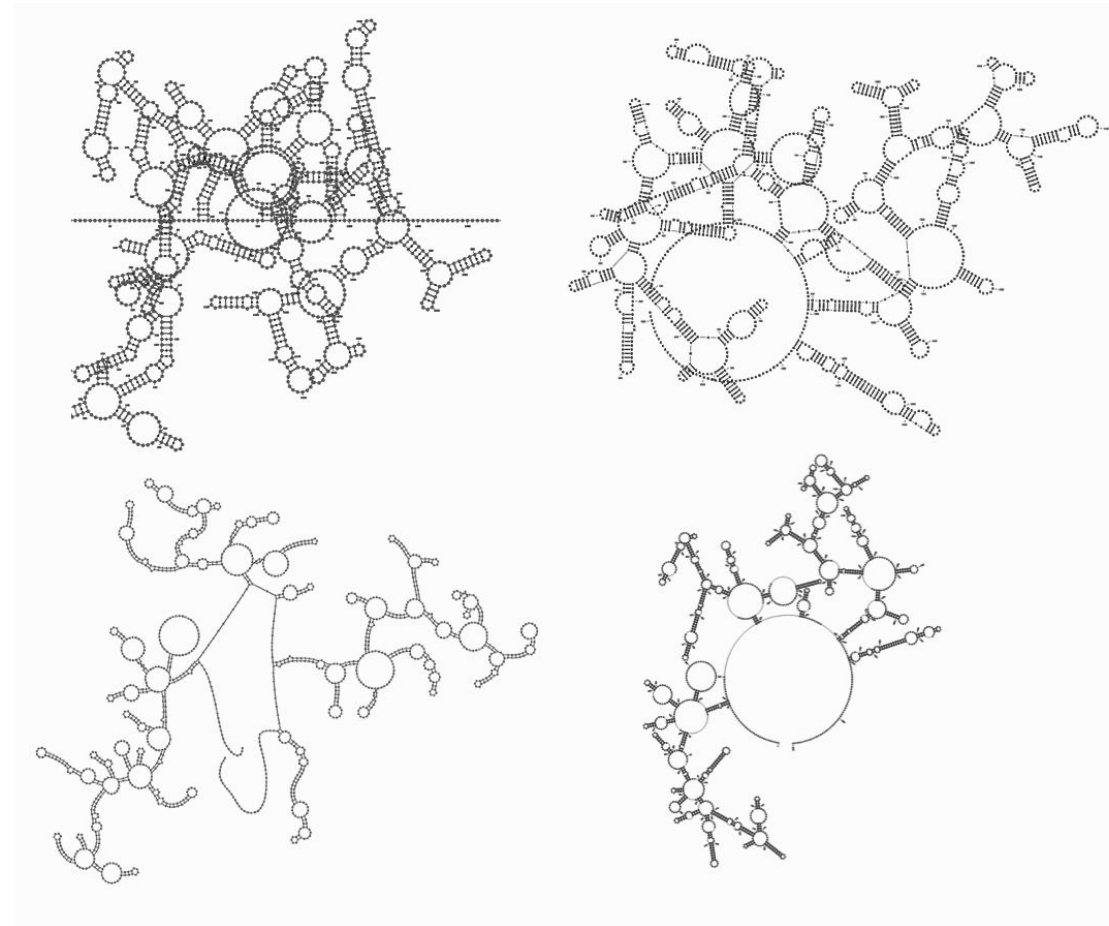


The goal is to display
RNA secondary structure in
consistent,
reproducible,
and recognisable
layouts

But automatically generated diagrams are not always useful



**Manually
curated layout**



Other RNA 2D visualisation software

Traveler software



David Hoksza

Elias and Hoksza *BMC Bioinformatics* (2017) 18:487
DOI 10.1186/s12859-017-1885-4

BMC Bioinformatics

SOFTWARE

Open Access

TRAVeLer: a tool for template-based RNA secondary structure visualization



Richard Elias and David Hoksza* 

Abstract

Background: Visualization of RNA secondary structures is a complex task, and, especially in the case of large RNA structures where the expected layout is largely habitual, the existing visualization tools often fail to produce suitable visualizations. This led us to the idea to use existing layouts as templates for the visualization of new RNAs similarly to how templates are used in homology-based structure prediction.

Results: This article introduces Traveler, a software tool enabling visualization of a target RNA secondary structure using an existing layout of a sufficiently similar RNA structure as a template. Traveler is based on an algorithm which converts the target and template structures into corresponding tree representations and utilizes tree edit distance coupled with layout modification operations to transform the template layout into the target one. Traveler thus accepts a pair of secondary structures and a template layout and outputs a layout for the target structure.

Conclusions: Traveler is a command-line open source tool able to quickly generate layouts for even the largest RNA structures in the presence of a sufficiently similar layout. It is available at <http://github.com/davidhoksza/traveler>.

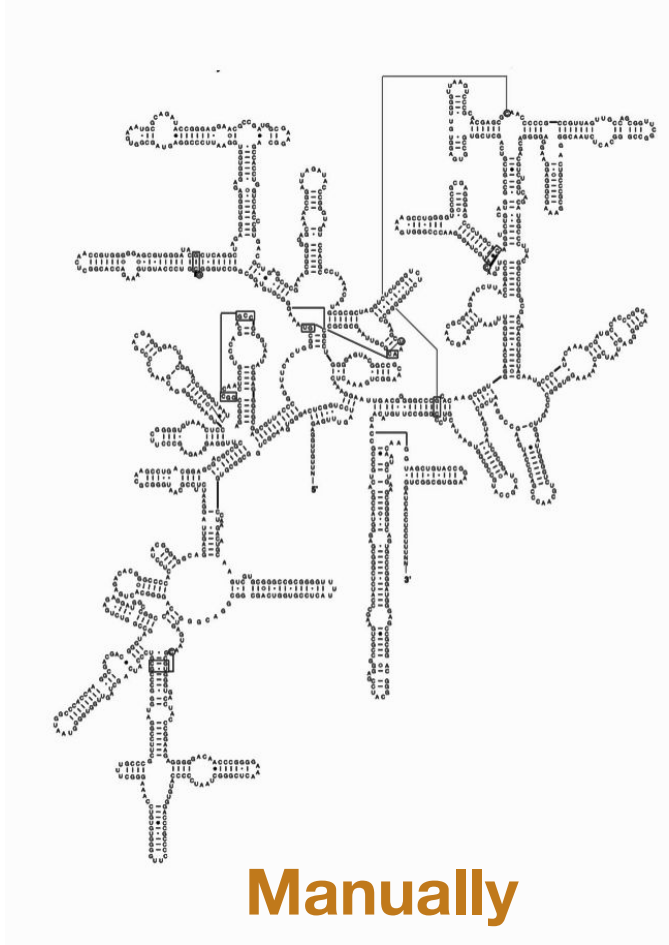
Keywords: Visualization, RNA secondary structure, Template-based modeling, Software tool

- Traveler visualises RNA structure for a given sequence using templates
- Is able to reposition the nucleotides to accommodate insertions

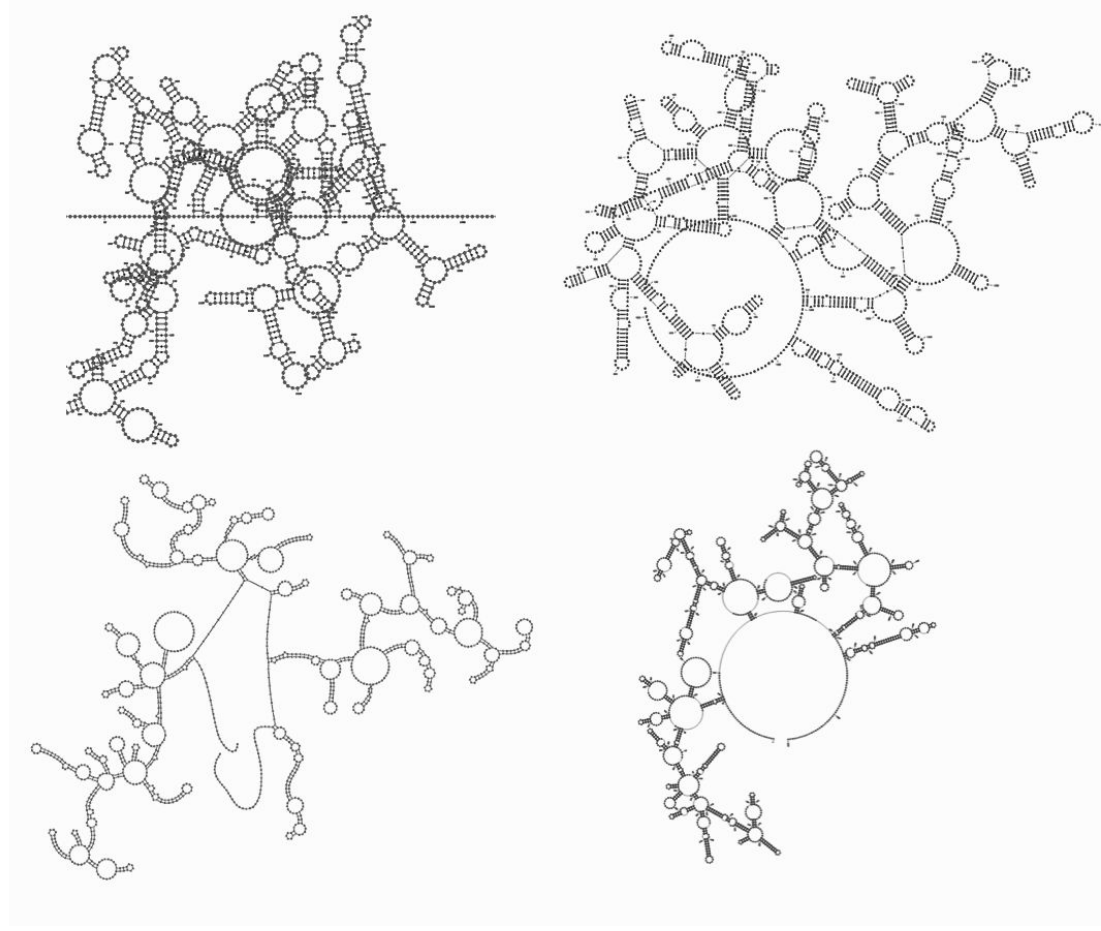
R2DT templates

- **A template must have:**
 - **Sequence**
 - **Secondary structure**
 - **(x, y) coordinates for each nucleotide**
- Traveler converts the target and template structures into corresponding tree representations and uses tree edit distance coupled with layout modification operations to transform the template layout into the target one
- Traveler can also use an Infernal alignment to map the layouts

With R2DT we are able to go from this 

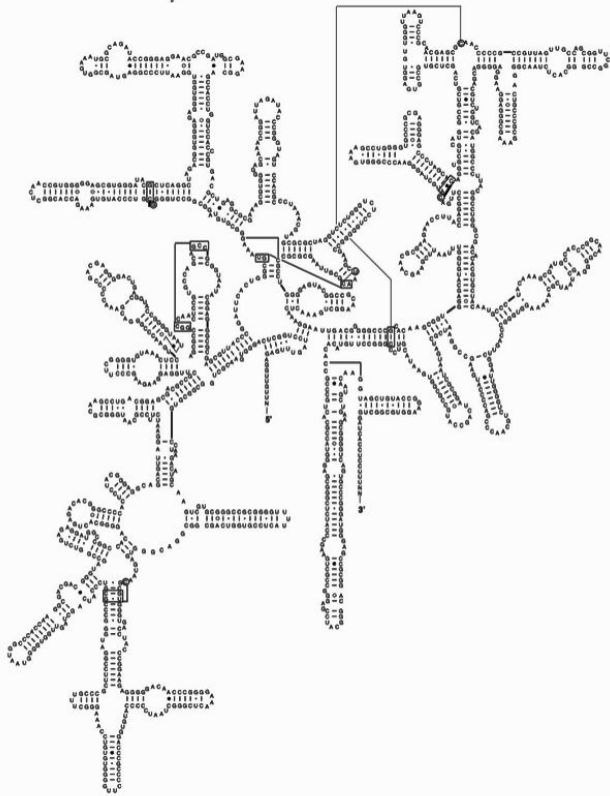


**Manually
curated layout**

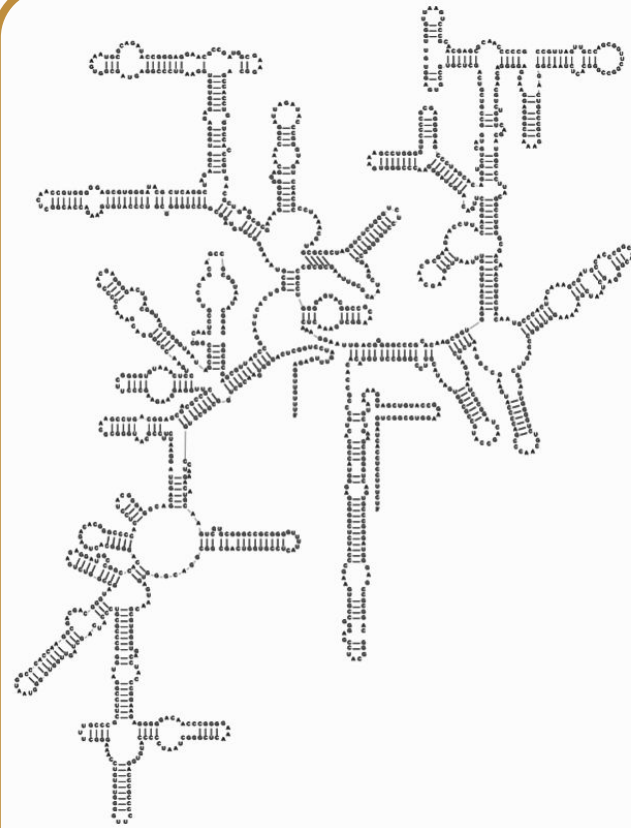


Other RNA 2D visualisation software

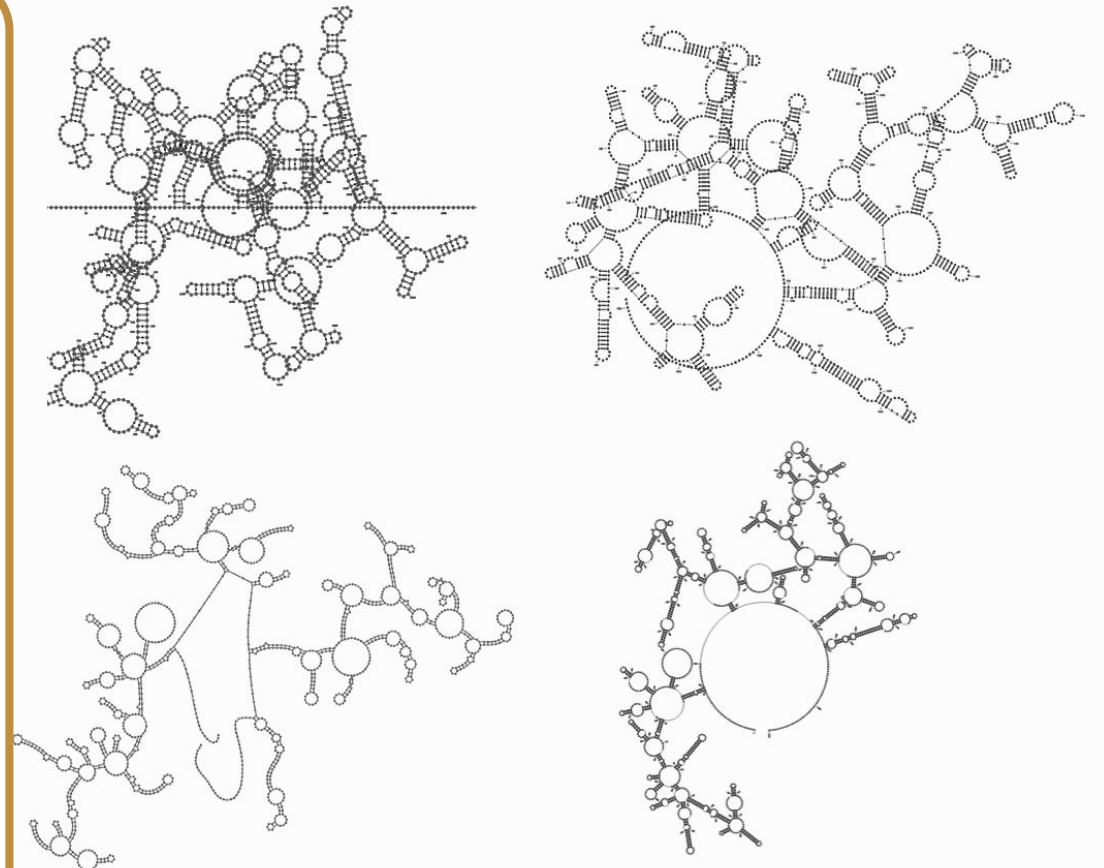
To this 📍



**Manually
curated layout**



R2DT



Other software

R2DT = RNA 2D Templates



This is where R2DT worked for the very first time in June 2018

RNA Society, Berkeley 2018

R2DT library includes >3,600 RNA 2D templates

RNA type	Template source	Number of templates	Manually curated?
SSU rRNA	CRW (covariation-based)	654	Yes
	RiboVision (3D-based)	8	Yes
LSU rRNA	RiboVision (3D-based)	21	Yes
5S rRNA	CRW	200	Yes
tRNA	GtRNAdb	74	Yes
RNase P	RNase P Database	17	Yes
		2 (3D-based)	Yes
Small RNAs	Rfam	2,671	No

R2DT release 1.0

**There are thousands of RNA families
and each needs a template.**

**How do we select which template to
use? And do it fast?**

Ribovore

- For each template, build an Infernal **covariance model** using the template sequence and secondary structure
- Rapidly select best-scoring template using Ribovore




Eric Nawrocki

SOFTWARE

Open Access

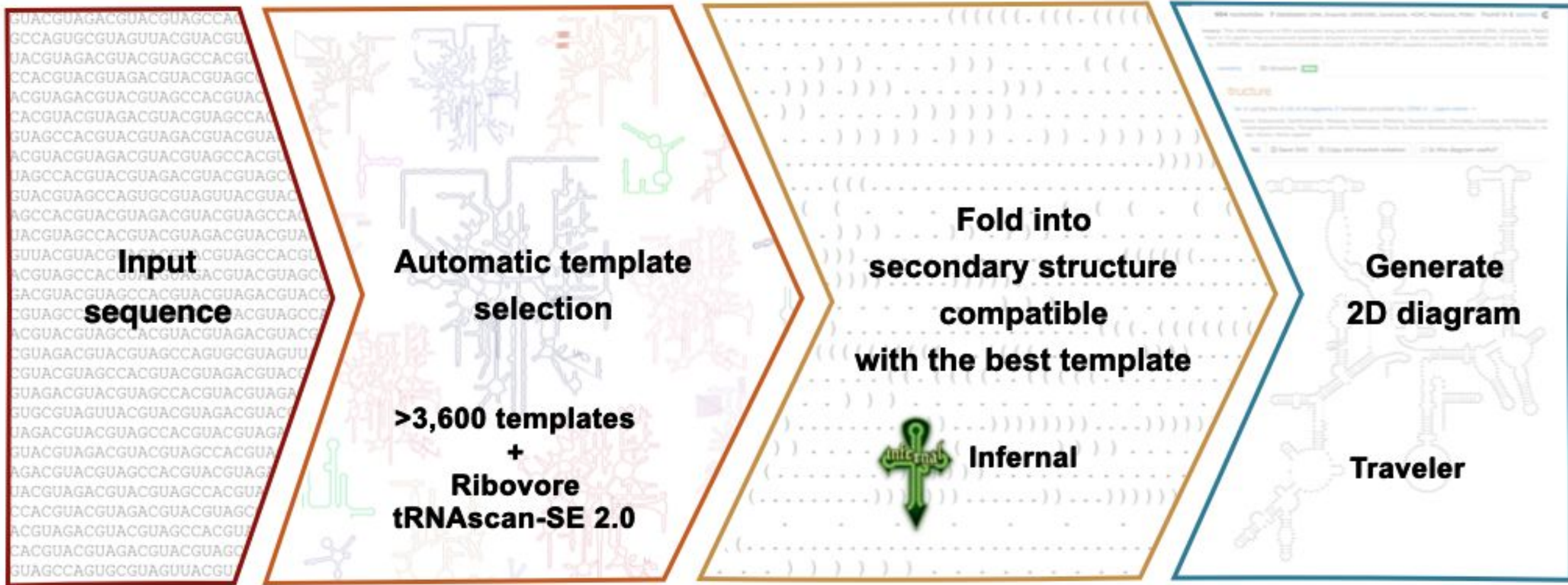
Ribovore: ribosomal RNA sequence analysis for GenBank submissions and database curation



Alejandro A. Schäffer^{1,2}, Richard McVeigh², Barbara Robbertse², Conrad L. Schoch², Anjanette Johnston², Beverly A. Underwood², Ilene Karsch-Mizrachi² and Eric P. Nawrocki^{2*} 

<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-021-04316-z>

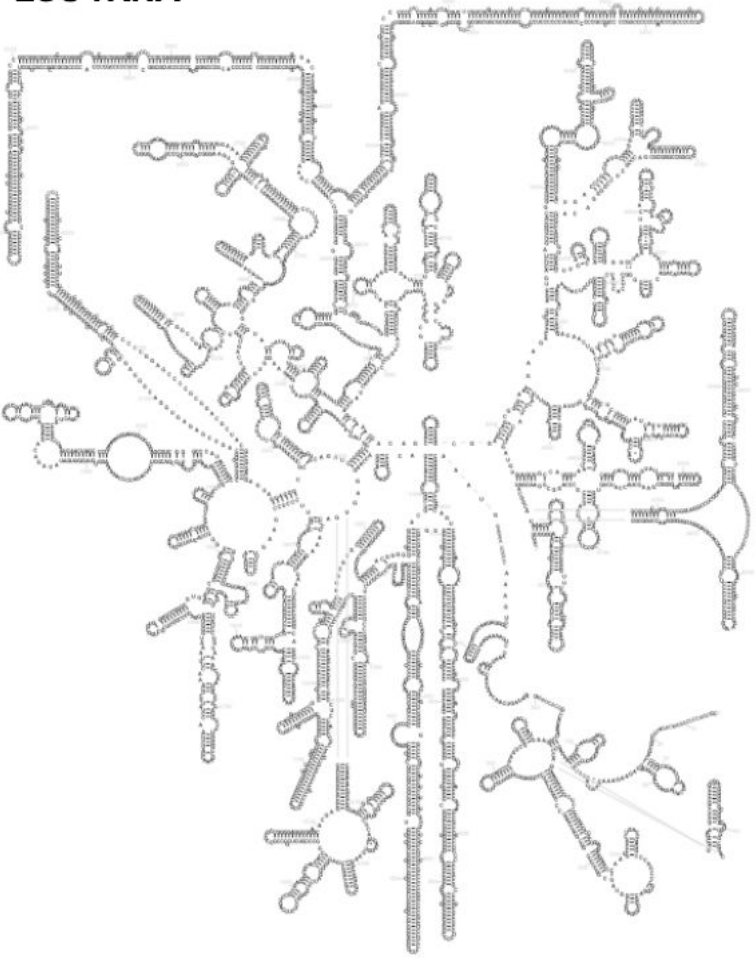
R2DT uses Ribovore, tRNAscan-SE, Infernal, and Traveler



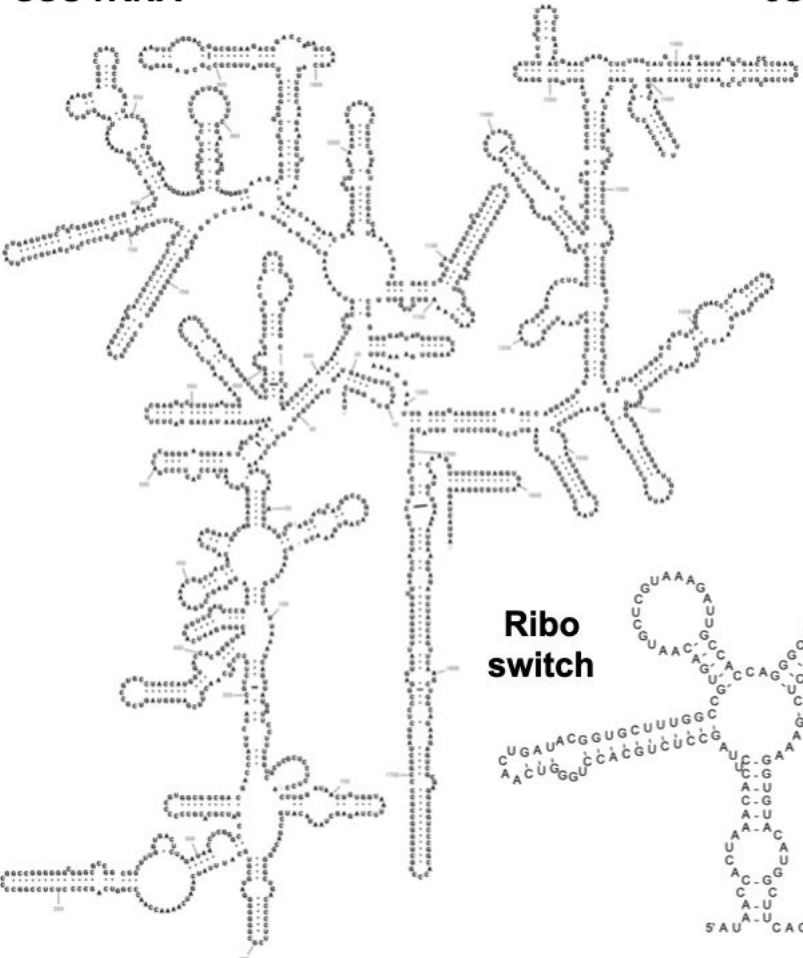


Examples

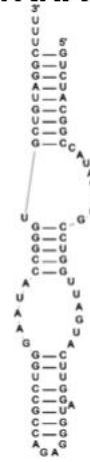
LSU rRNA



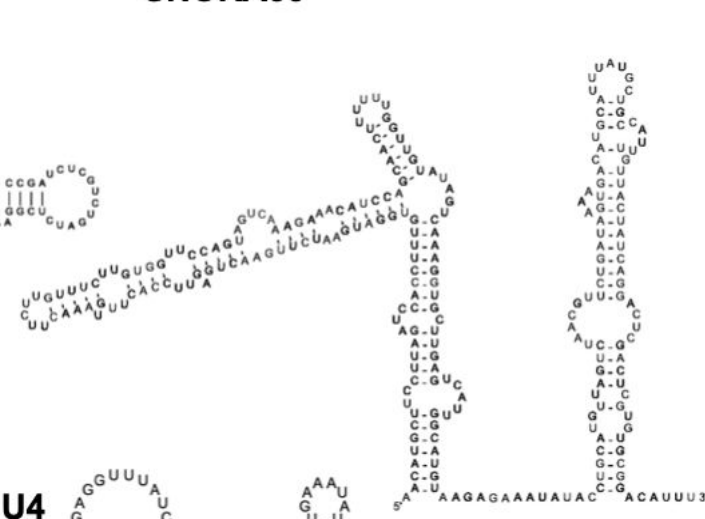
SSU rRNA



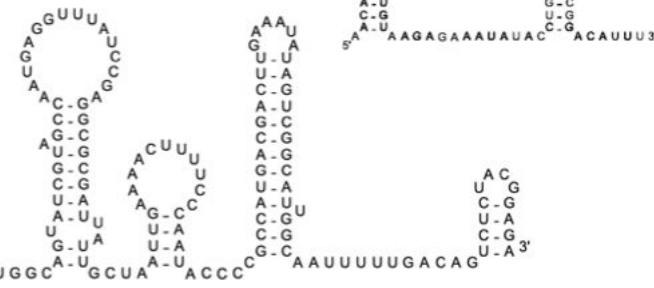
5S rRNA



SNORA53



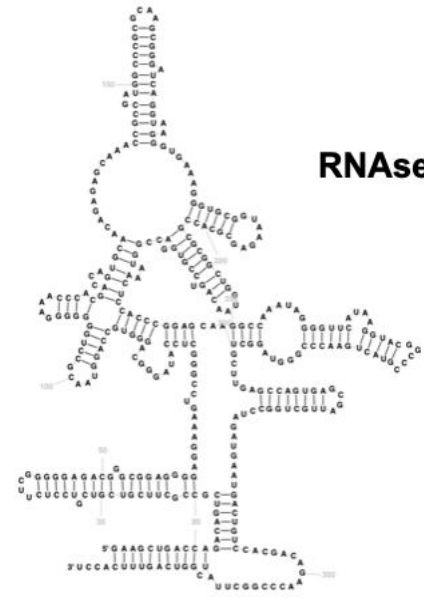
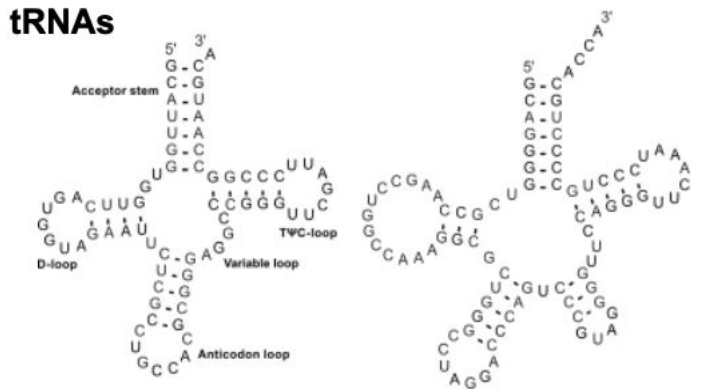
U4



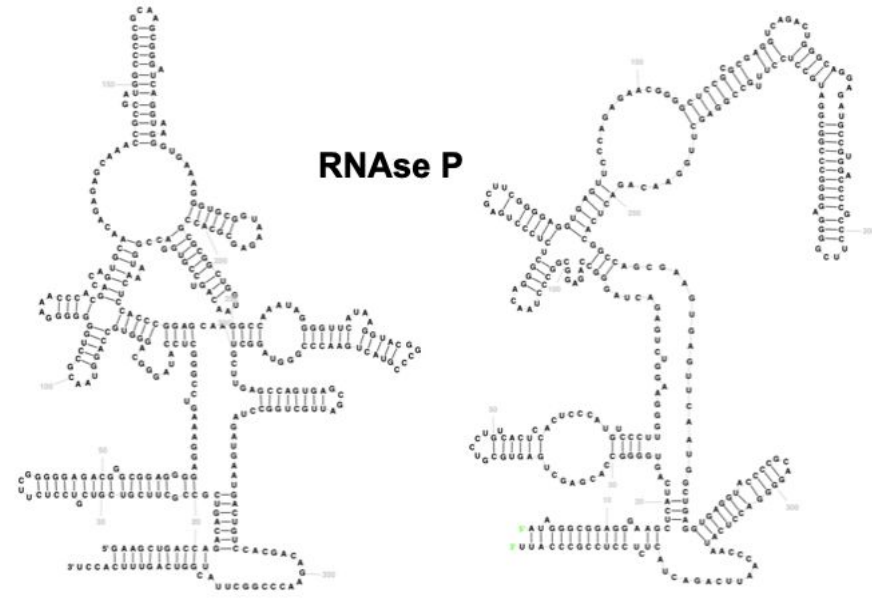
Ribo switch



tRNAs



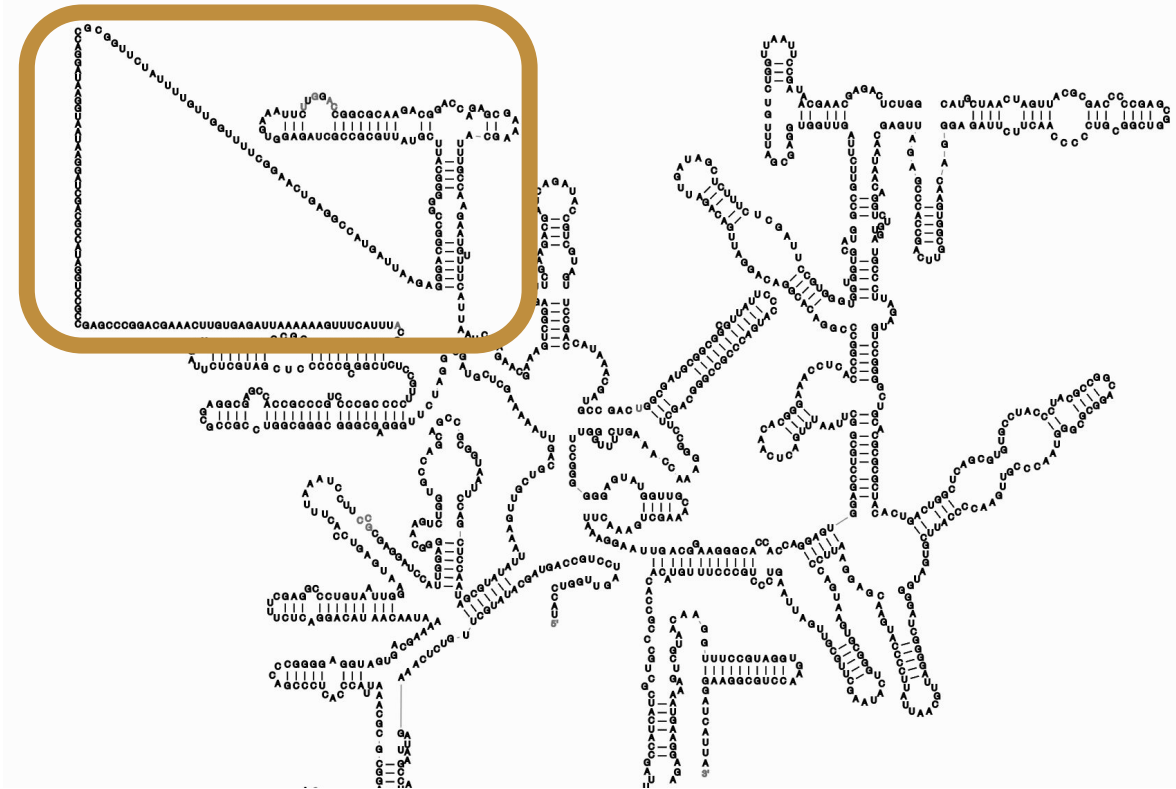
RNase P



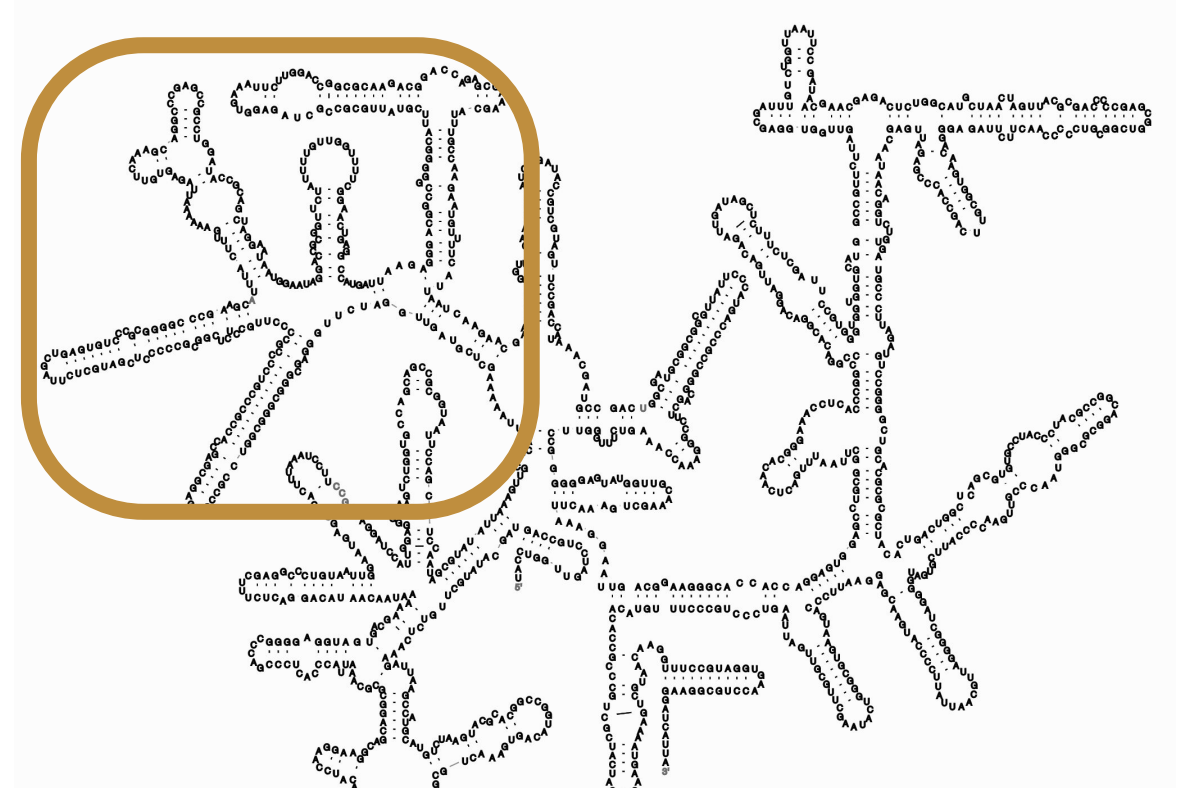
3D structure based templates from COOL



Anton S. Petrov



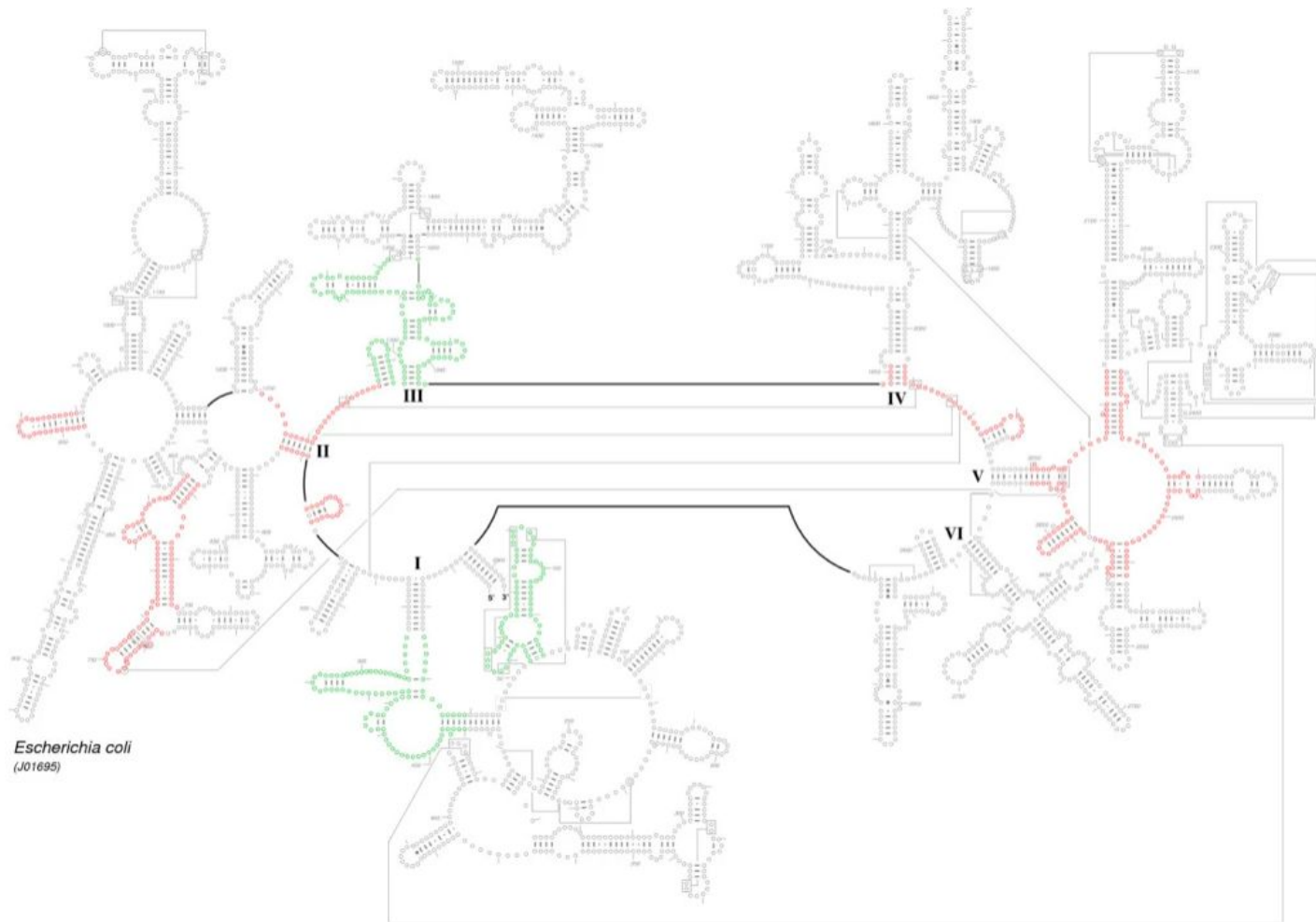
CRW covariation-based template



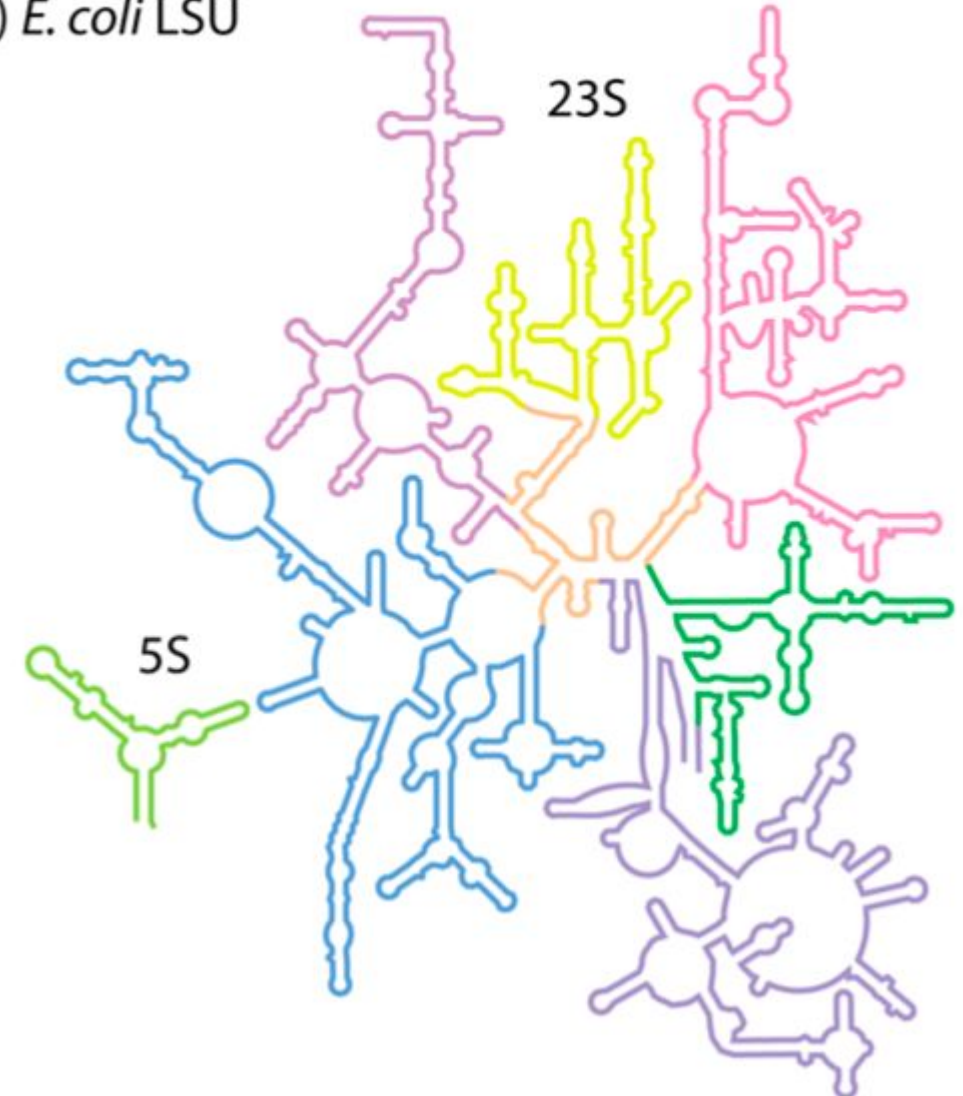
3D-based template

R2DT uses single page LSU layouts

Anton S. Petrov



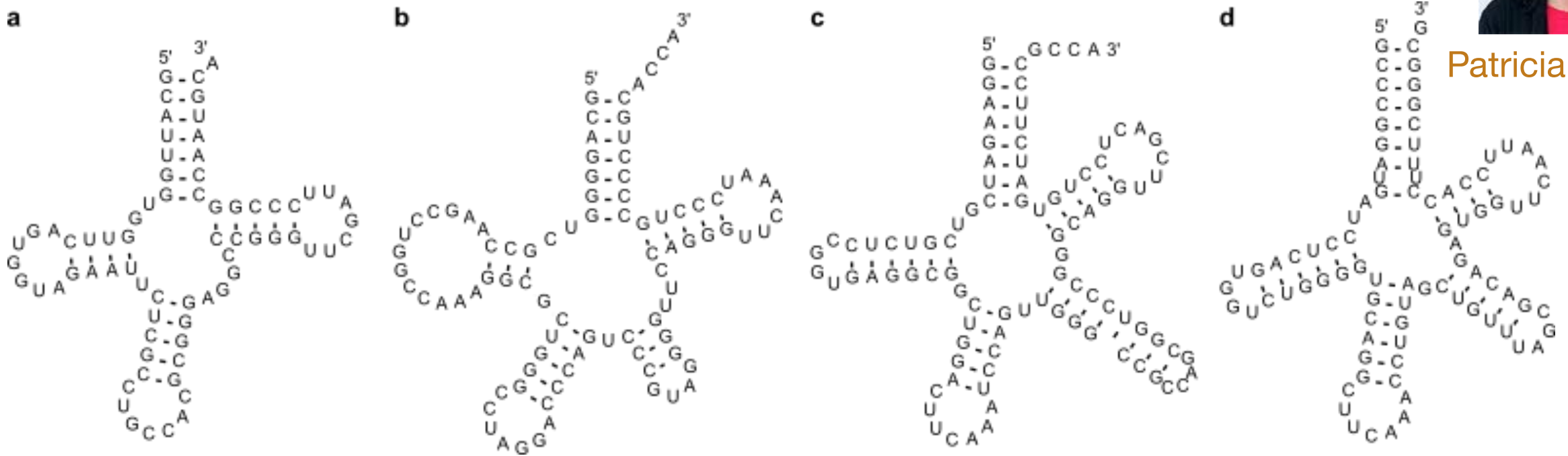
a) *E. coli* LSU



tRNA templates



Patricia Chan



- **74 isotype-specific templates** for Bacteria, Eukaryotes, and Archaea
- Generic Type I and Type II templates for pseudogenes and organellar tRNAs

**For the rest of RNA families
R2DT uses Rfam consensus
secondary structures
generated with R2R**

Thanks to Zasha Weinberg!



As with any prediction method, R2DT results need to be used with caution

Colour coding the alignment of the sequence to the template

- **Insertions** are shown in red
- **Mutations** are shown in magenta
- Nucleotides that required **repositioning** are shown in blue
- Nucleotides that are **identical** to the template are shown in black











ARTICLE



<https://doi.org/10.1038/s41467-021-23555-5>

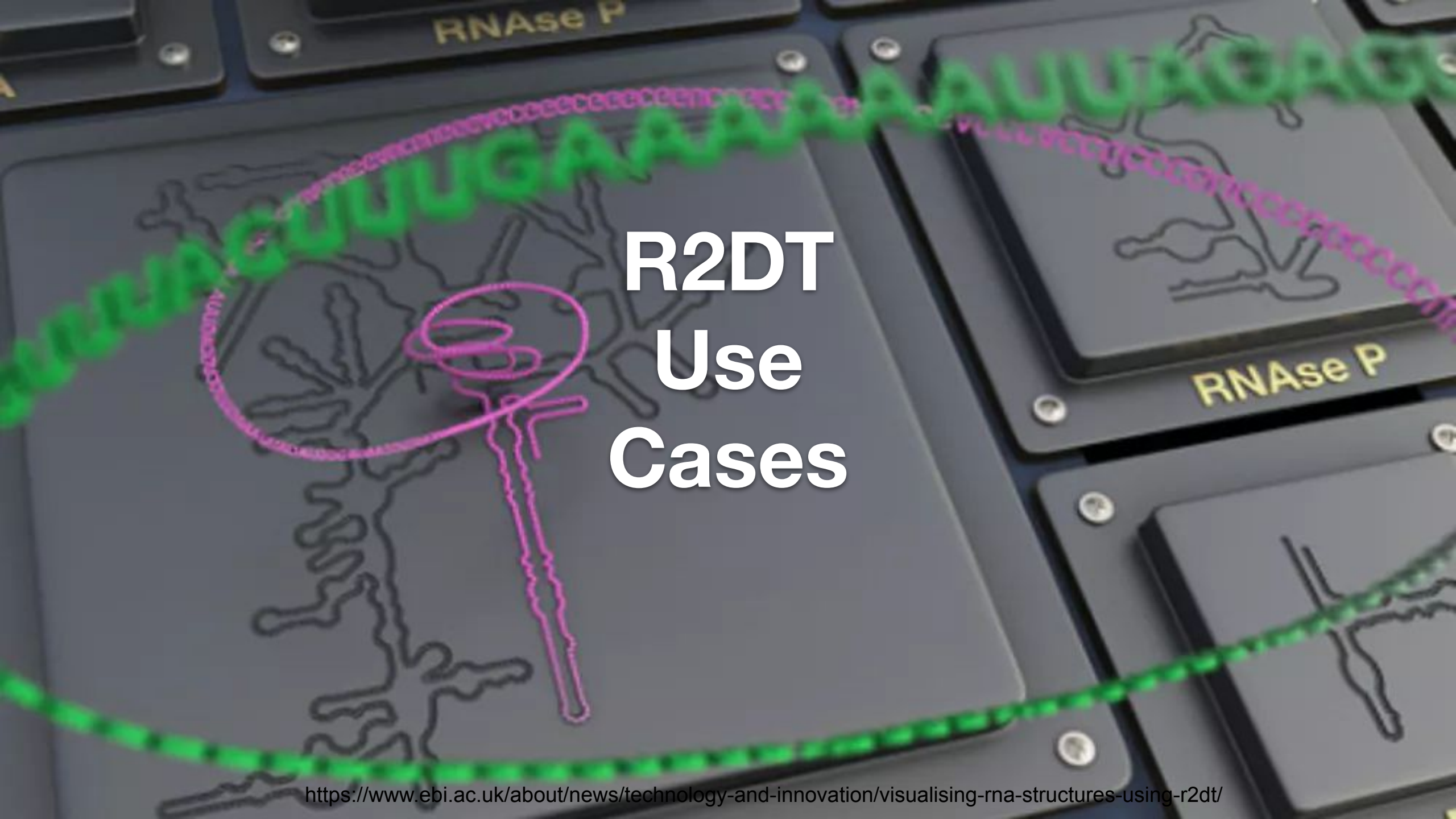
OPEN

R2DT is a framework for predicting and visualising RNA secondary structure using templates

Blake A. Sweeney^{1,7}, David Hoksza ^{2,7}, Eric P. Nawrocki³, Carlos Eduardo Ribas ¹, Fábio Madeira ¹,
Jamie J. Cannone⁴, Robin Gutell⁴, Aparna Maddala⁵, Caeden D. Meade⁵, Loren Dean Williams ⁵,
Anton S. Petrov ⁵, Patricia P. Chan ⁶, Todd M. Lowe⁶, Robert D. Finn ^{1,8} & Anton I. Petrov ^{1,8}✉

<https://www.nature.com/articles/s41467-021-23555-5>

<https://github.com/rnacentral/r2dt>



R2DT
Use
Cases

R2DT is integrated into the RNACentral sequence search

Sequence search

Local alignment using nhmmer

AAGACUAUACUUUCAGGGAUCAUUUCUAUAGUGUGUUACUAGAGAAGUUUCUCUGAACGUGUAGAGCACCGAAAAC
 CACGAGGAAGAGAGGAGCGUUUCUCCUGAGCGUGAAGCCGGCUUUCUGGCGUUGCUUUGGCUGCAACUGCCGUC
 AGCCAUUGAUGAUCGUUCUUCUCCGUAUUGGGGAGUGAGAGGGAGAGAACGCGGUCUGAGUGGU

Search **Clear** **Upload file**
 Up to 50 queries (beta)

Examples: miRNA hsa-let-7a-1 SNORD3A lysine riboswitch 5S rRNA 16S rRNA NKILA lncRNA

Identical match: Homo sapiens small nucleolar RNA, C/D box 3A (SNORD3A)

Secondary structure

R2DT
 Visualise RNA secondary structure in standard orientations using RNA 2D Templates.

View

Rfam classification

Family	Accession	Start	End	Bit score	E-value	Strand	Alignment
Small nucleolar RNA U3	RF00012	1	217	179.9	1.8e-44	+	Show

Similar sequences 986 of 10023

Text search within results **Filter** **Clear** **Sort by E-value (min to max) - default** **Hide alignments** **See details** **Download**

RNA types

- snoRNA (925)
- lncRNA (42)
- snRNA (17)
- other (1)
- sRNA (1)

Organisms

- Homo sapiens (63)
- Mus musculus (22)

Identical match: Homo sapiens small nucleolar RNA, C/D box 3A (SNORD3A)

URS000053962A_9606 **ENA** **CCDS** **GeneCards** **HGNC** **LNCipedia** **RefSeq** **Rfam** **snoDB** **snoPY**

Query 1 AAGACUAUACUUUCAGGGAUCAUUUCUAUAGUGUGUUACUAGAGAAGUUUCUCUGAACGUGUAGAGCACCGAAAACACGAGGAAGAGAGG 93
 Sbjct 1 AAGACUAUACUUUCAGGGAUCAUUUCUAUAGUGUGUUACUAGAGAAGUUUCUGAAGCGUGUAGAGCACCGAAAACACGAGGAAGAGAGG 93

Query 92 UAGCGUUUUUCUCCUGAGCGUGAAGCCGGCUUUCUGGCGUUGCUUUGGCGCAACUGCCGUCAGCCAUUGAUGAUCGUUCUUCUCCGUAUU 18
 Sbjct 92 UAGCGUUUUUCUCCUGAGCGUGAAGCCGGCUUUCUGGCGUUGCUUUGGCGCAACUGCCGUCAGCCAUUGAUGAUCGUUCUUCUCCGUAUU 18

Homo sapiens (human) small nucleolar RNA, C/D box 3A (SNORD3A)

Rfam HOME SEARCH BROWSE FTP BLOG HELP **CONTACT** Search Rfam

Search Rfam

Sequence search

Powered by RNACentral | Local alignment using nhmmer Job id: cf4857-892-4715-9892-492030300

GGACGAGGCCCGCCGAGUAGGUGUCCUACAGGGGAGUUCGGGACGGGCGAAGAGCGAGGCGCCGAGCGAGCGAGUUCUCC
 GCUCUCUUGGCGGGGUGAAGUACUCCUACUCCUCCGAAAGCGGAGCCGUCCA

Examples: lysine riboswitch 16S SNORD3A

Identical match: Lysine Riboswitch RNA from PDB 3DOU, chain A and 1 other sequence

Rfam classification

Family	Accession	Start	End	Bit score	E-value	Strand
Lysine riboswitch	RF00168	1	160	86.2	4.7e-23	+

Secondary structure

R2DT
 Visualise RNA secondary structure in standard orientations using RNA 2D Templates.

View

Text search within results **Filter** **Clear** **Sort by E-value (min to max) - default** **Hide alignments** **See details** **Download**

RNA types

- Lysine riboswitch RNA from Thermotoga maritima (PDB 4ER), chain A
- tRNA (411)
- other (251)
- misc RNA (37)
- sRNA (31)
- rRNA (14)

Rfam

Home Search Help Links About

GIRNADB
 Prediction of conserved genes

Search by Sequence

Choose search method **BLAST Search** **RNACentral Search**

Powered by RNACentral | Local alignment using nhmmer Job id: cf926fb-75d8-4493-b865-a55aded9e47

GGGGUUAJAGCUCAGUGGAGAGCAUUGACUGCAGAGCAAGAGTCCCGGUCAAAU
 CCGGGUCCCCU

Search **Clear** **Upload file**
 Up to 50 queries (beta)

Examples: tRNA-Cys-GCA-2-2 in human tRNA-Met-CAT-1-1 in Methanocaldococcus jannaschii

Identical match: Homo sapiens tRNA-Cys (anticodon GCA) 2-1 (TRC-GCA2 1 to 4)

Secondary structure

R2DT
 Visualise RNA secondary structure in standard orientations using RNA 2D Templates.

View

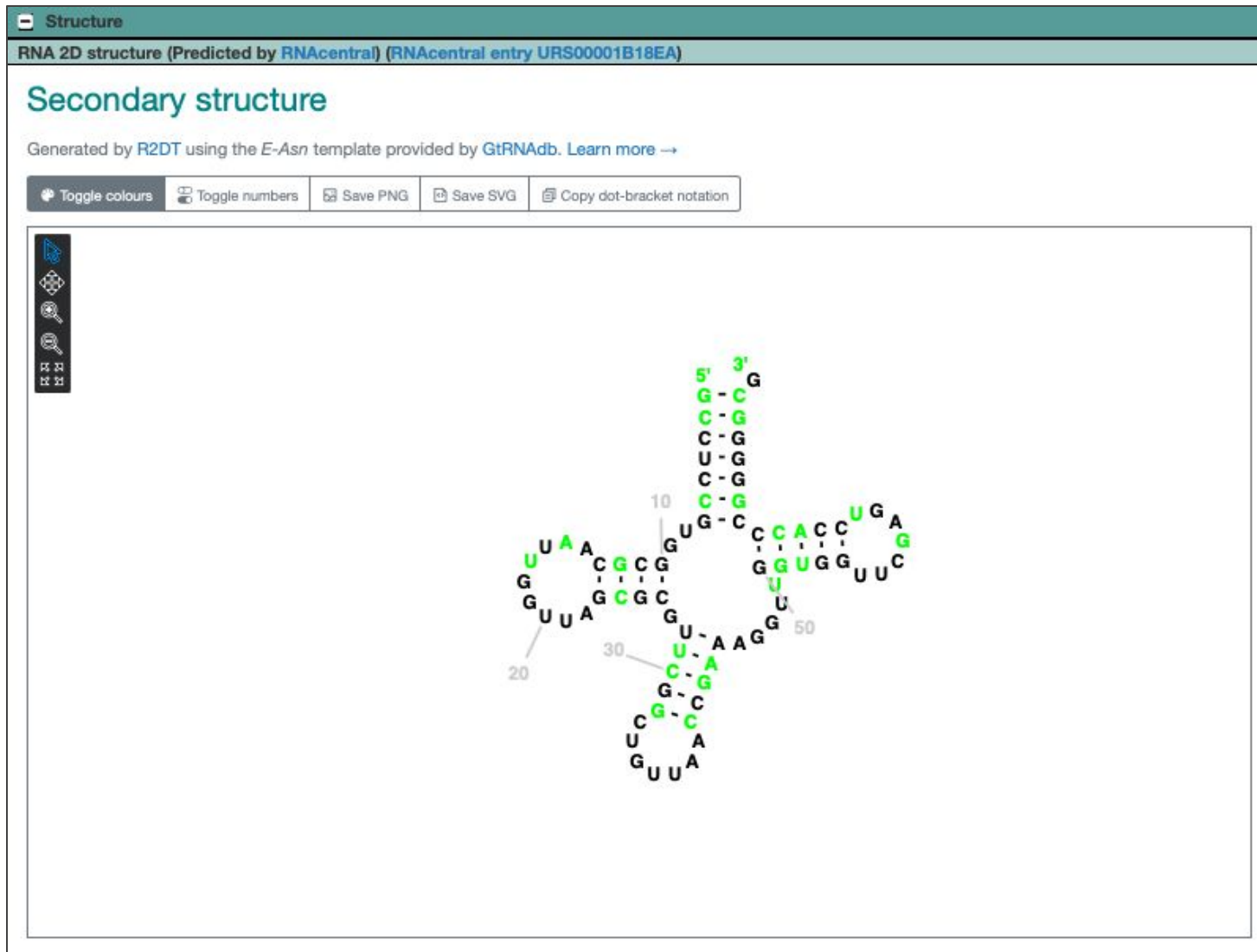
Similar sequences 1000 of 23406

Text search within **Filter** **Clear** **Sort by E-value (min to max) - r** **Hide alignments** **See details** **Download**

GtRNADB



R2DT is integrated into FlyBase - SGD is coming soon



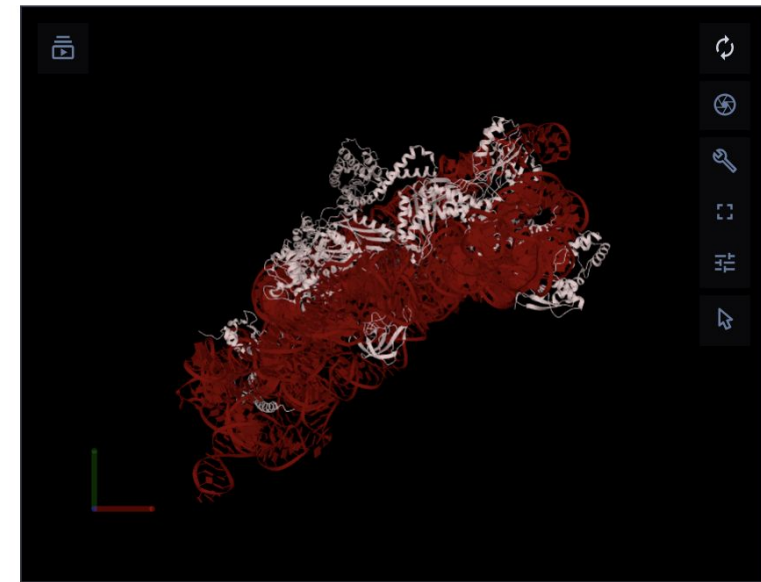
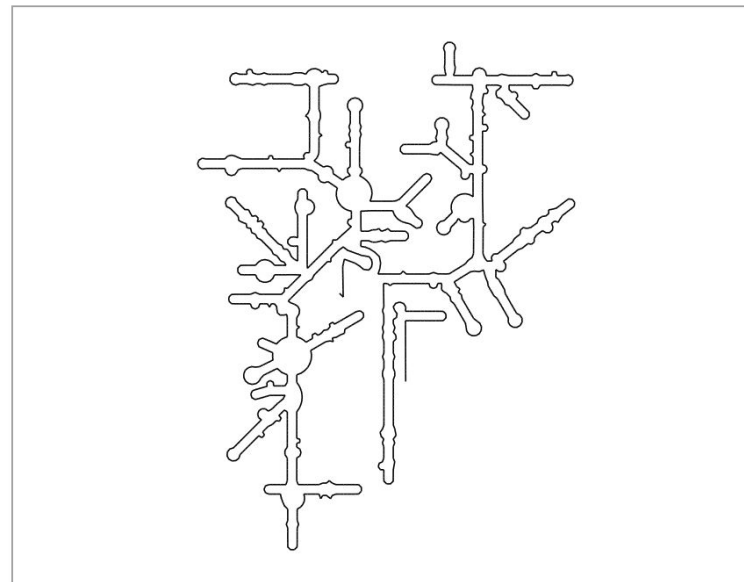
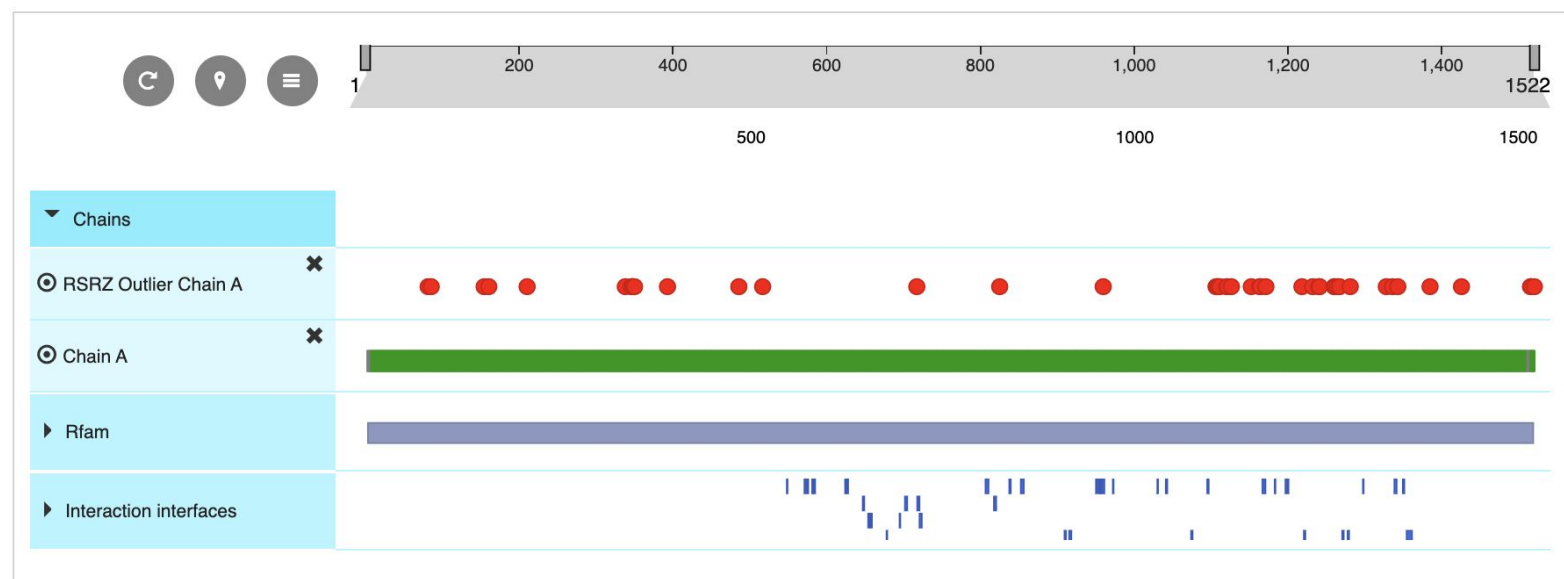
R2DT is integrated into PDBe

- Interactive 1D-2D-3D visualisation
- Clicking RNA 2D structure highlights the region in 3D and 1D

Thermus thermophilus small subunit ribosomal RNA (1J5E:A)

Problem: no 3D-derived base pairing information

▼ Visualisation



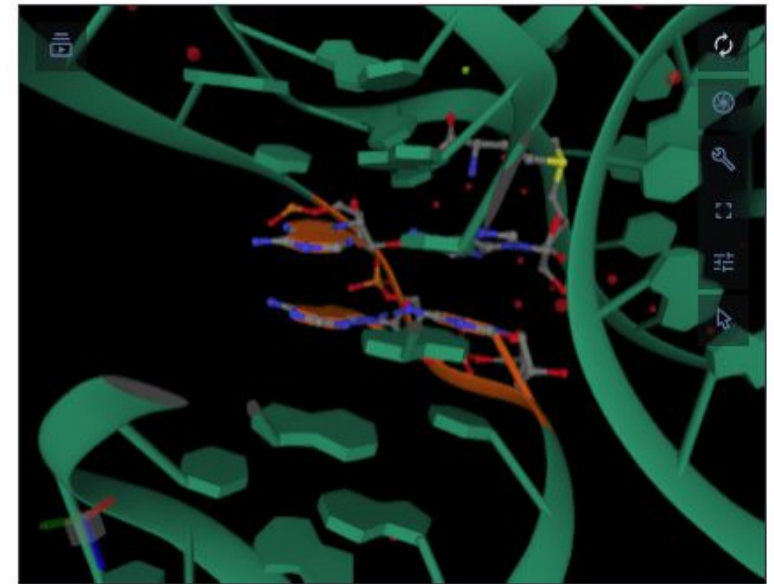
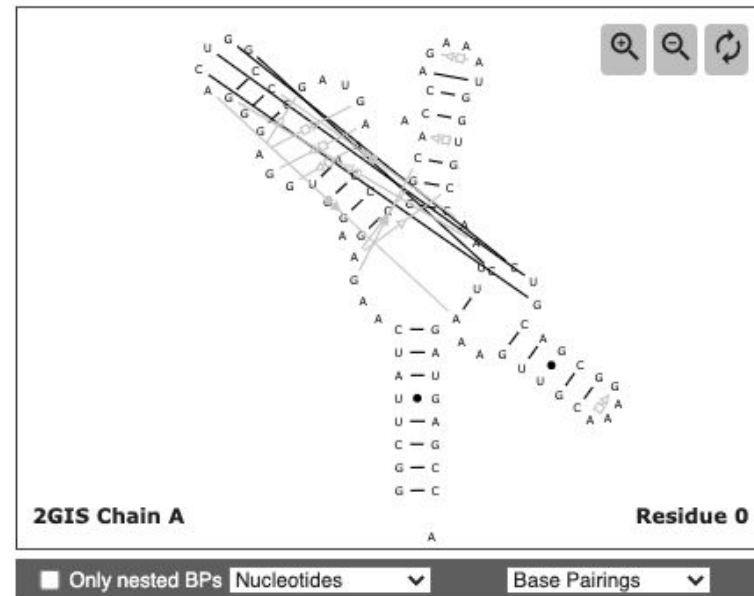
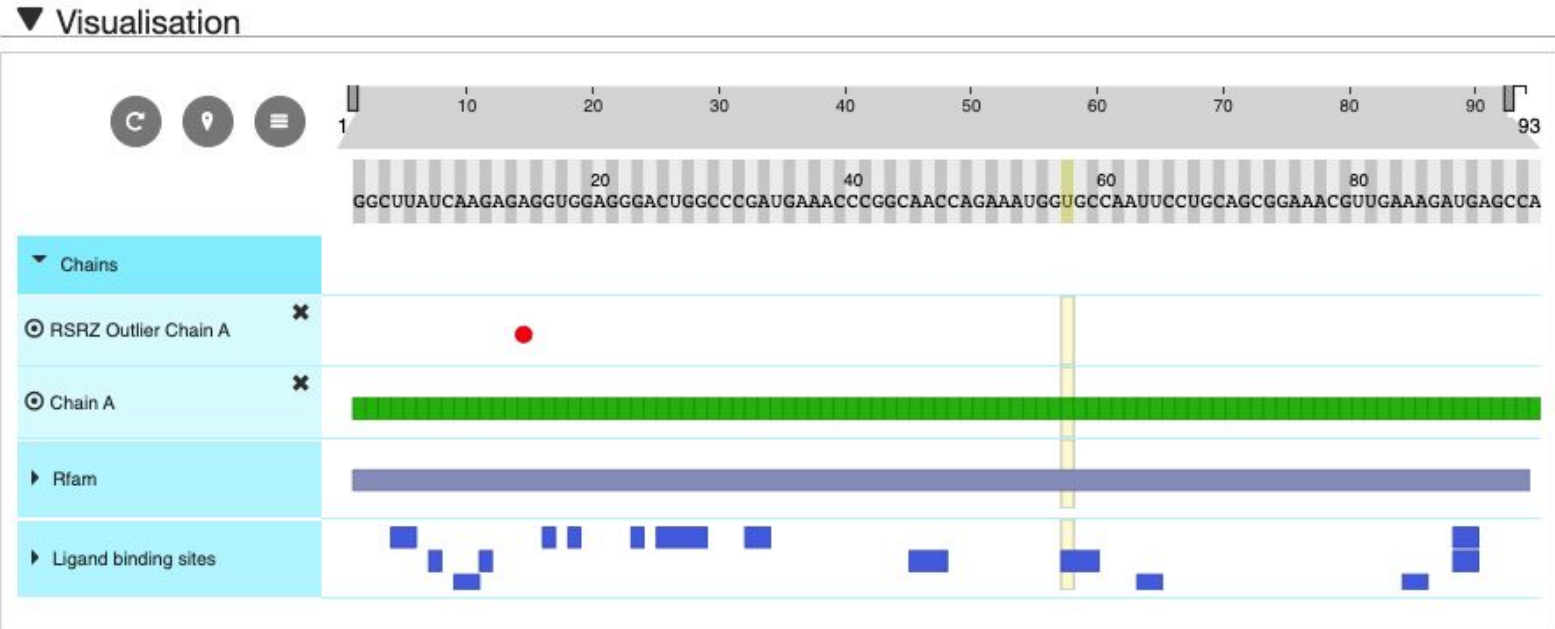
New PDBe widget

- Base pair annotations
- Leontis-Westhof symbols
- Both nested and long range interactions



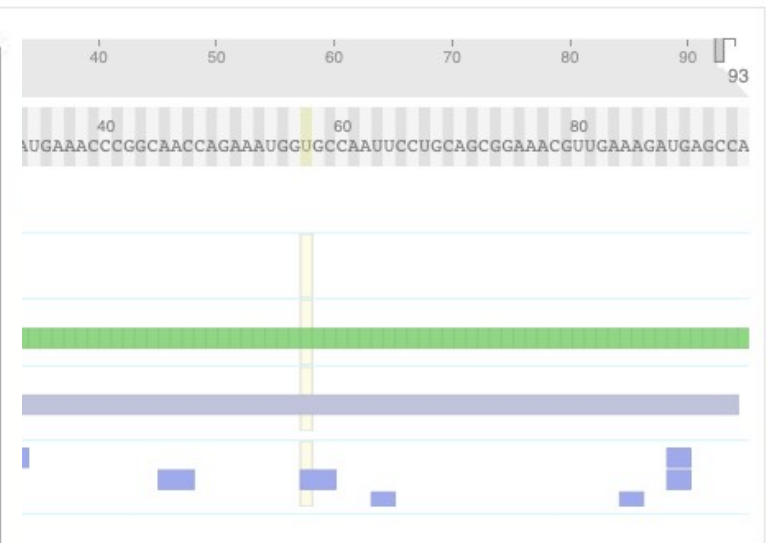
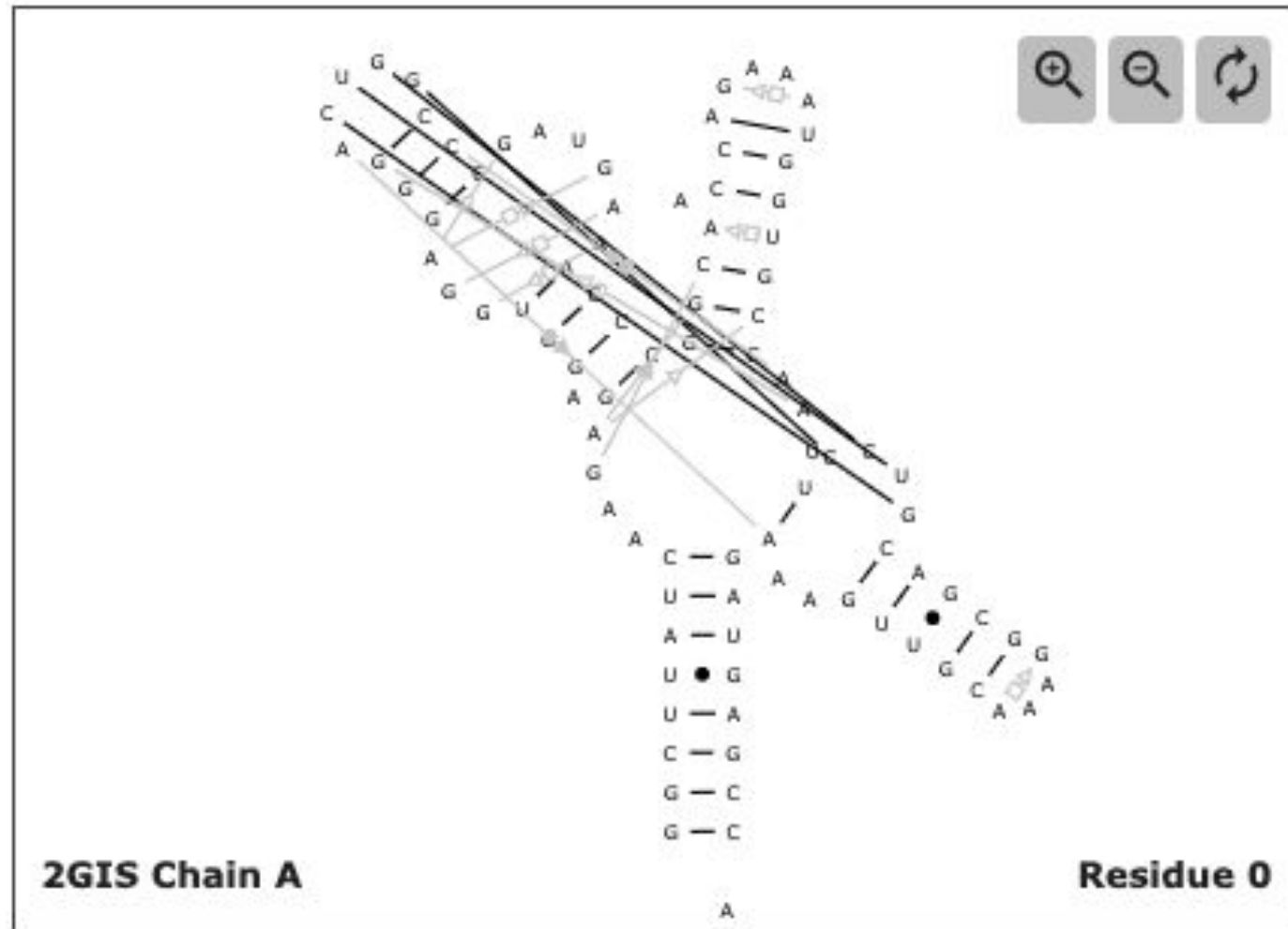
Holly McCann

<https://wwwdev.ebi.ac.uk/pdbe/entry/pdb/2gis/RNA/1>



R2DT displays non-canonical basepairs annotated by FR3D

▼ Visualisation



Live demo



<https://wwwdev.ebi.ac.uk/pdbe/entry/pdb/2gis/RNA/15FJC>

A young child wearing a cap and overalls stands on a wide set of stone steps. The steps lead up to a wall made of large, light-colored stone blocks. A sharp shadow of the wall is cast onto the steps from the right. The text "What's next?" is overlaid on the right side of the image.

**What's
next?**



New PDBe API serving FR3D basepairs

- Uses FR3D Python github.com/BGSU-RNA/fr3d-python
- Annotates basepairs for modified nucleotides
- Supports mmCIF files
- Currently under development - Stay tuned!



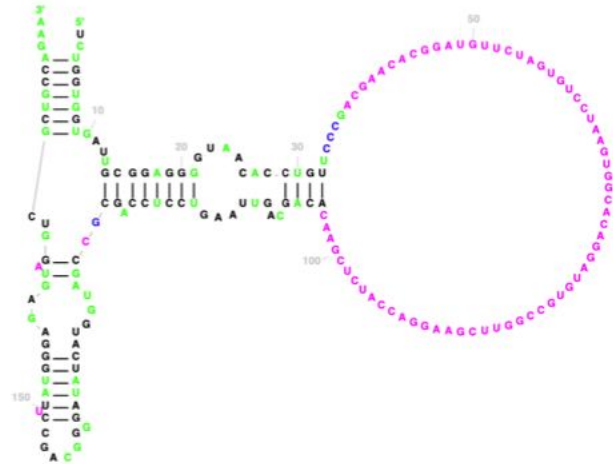
Craig Zirbel



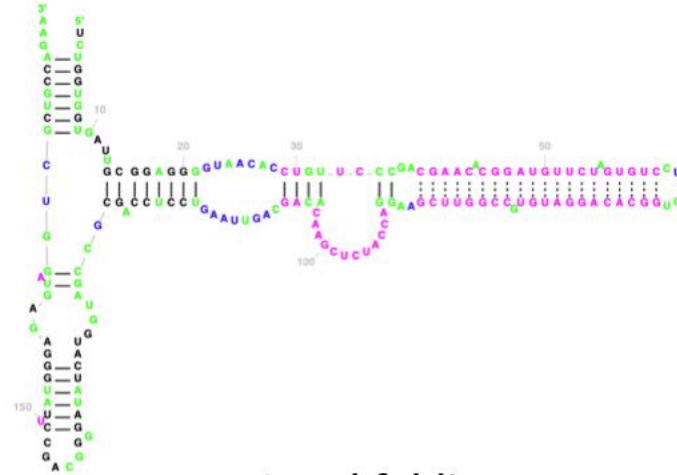
Constrained folding using RNAfold



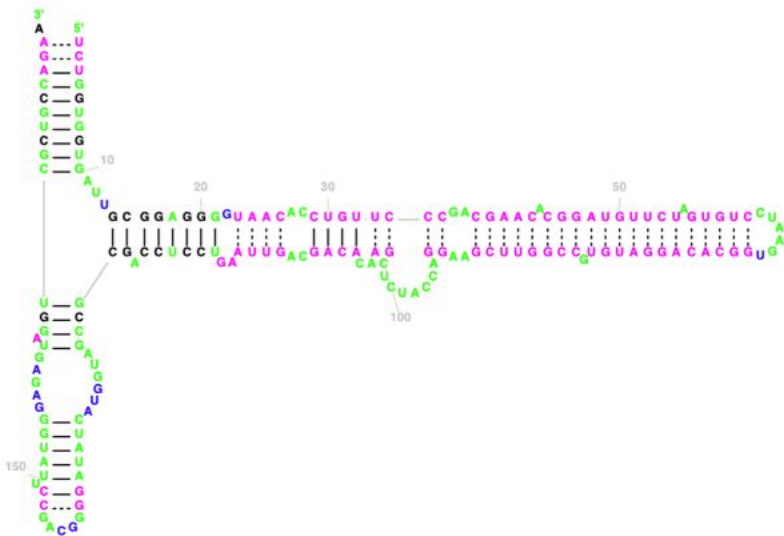
Holly McCann



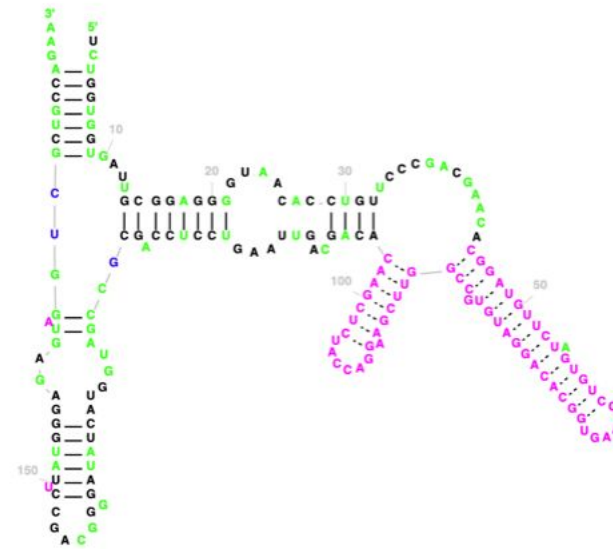
Original R2DT



Local folding



Global folding

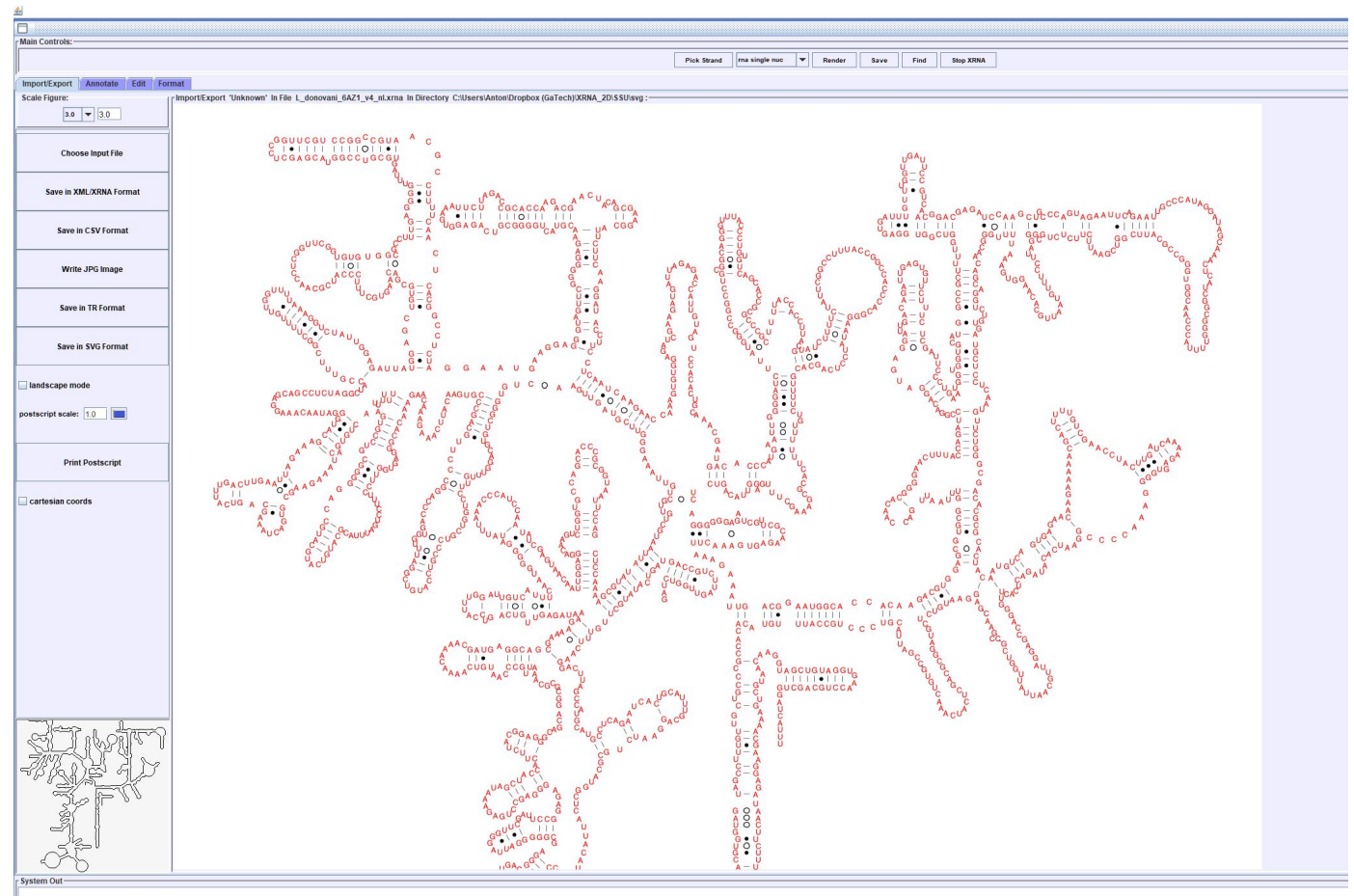


Global folding w/ss nucleotides enforced

R2DT can be used to produce new templates using XRNA-GT

1. Download R2DT results
2. Open in XRNA-GT
3. Manually edit the layout
4. Export in R2DT format
5. Submit as a new template

github.com/LDWLab/XRNA-GT





Interactive web-based editor

1. Export a diagram from R2DT
2. Edit it in a browser
3. Submit to R2DT as a new template

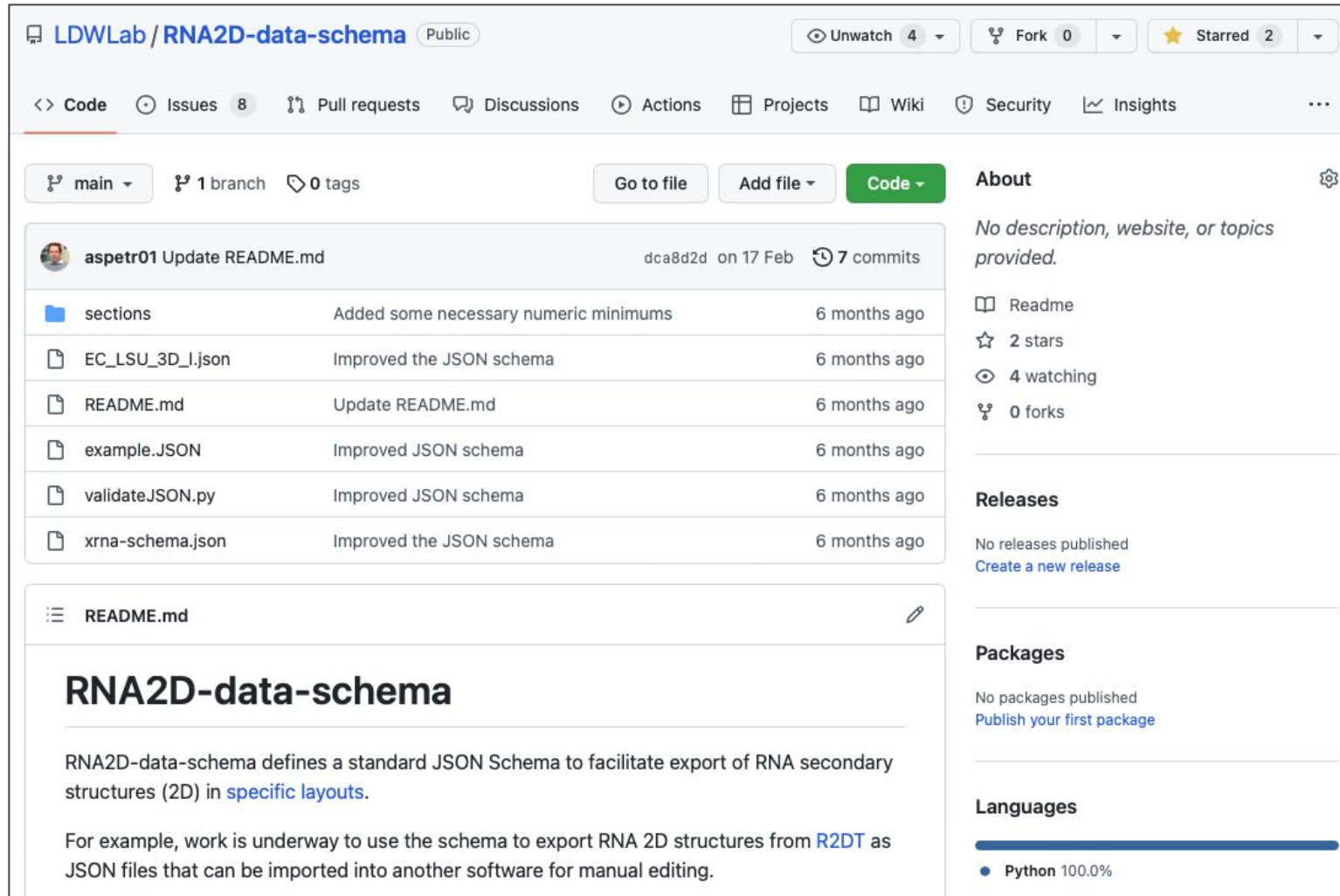


Caeden Meade

RNA 2D JSON Schema



Caeden Meade



LDWLab / RNA2D-data-schema Public

Unwatch 4 Fork 0 Starred 2

<> Code Issues 8 Pull requests Discussions Actions Projects Wiki Security Insights

main 1 branch 0 tags

Go to file Add file Code

Commit	Message	Time
aspetr01	Update README.md	dca8d2d on 17 Feb 7 commits
sections	Added some necessary numeric minimums	6 months ago
EC_LSU_3D_I.json	Improved the JSON schema	6 months ago
README.md	Update README.md	6 months ago
example.JSON	Improved JSON schema	6 months ago
validateJSON.py	Improved JSON schema	6 months ago
xrna-schema.json	Improved the JSON schema	6 months ago

README.md

RNA2D-data-schema

RNA2D-data-schema defines a standard JSON Schema to facilitate export of RNA secondary structures (2D) in [specific layouts](#).

For example, work is underway to use the schema to export RNA 2D structures from [R2DT](#) as JSON files that can be imported into another software for manual editing.

About
No description, website, or topics provided.

Readme
2 stars
4 watching
0 forks

Releases
No releases published
[Create a new release](#)

Packages
No packages published
[Publish your first package](#)

Languages

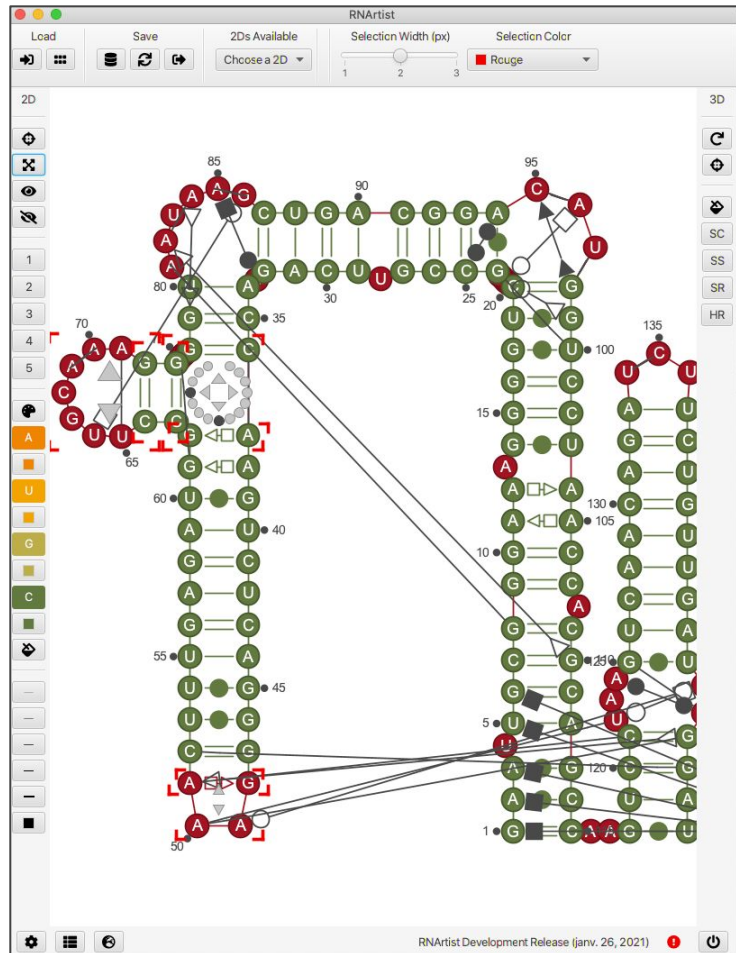
- Python 100.0%

Includes:

- Sequence and secondary structure
- (x,y) coordinates
- Colours, fonts, labels, additional annotations

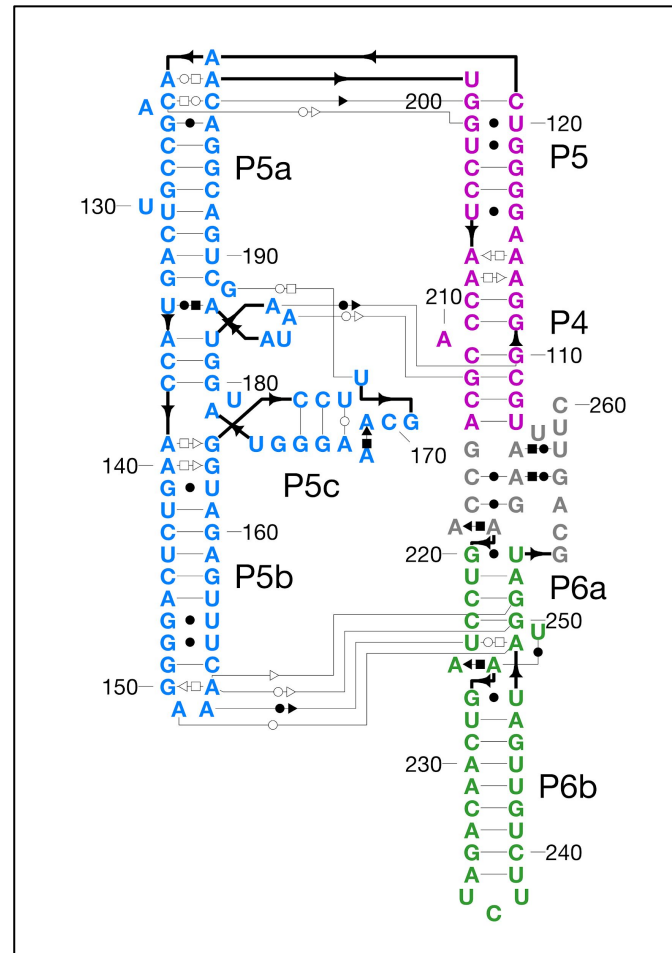
Can we make RNA visualisation software interoperable?

RNArtist



github.com/fjossinet/RNArtist

RiboDraw



github.com/ribokit/RiboDraw

Long term goal:

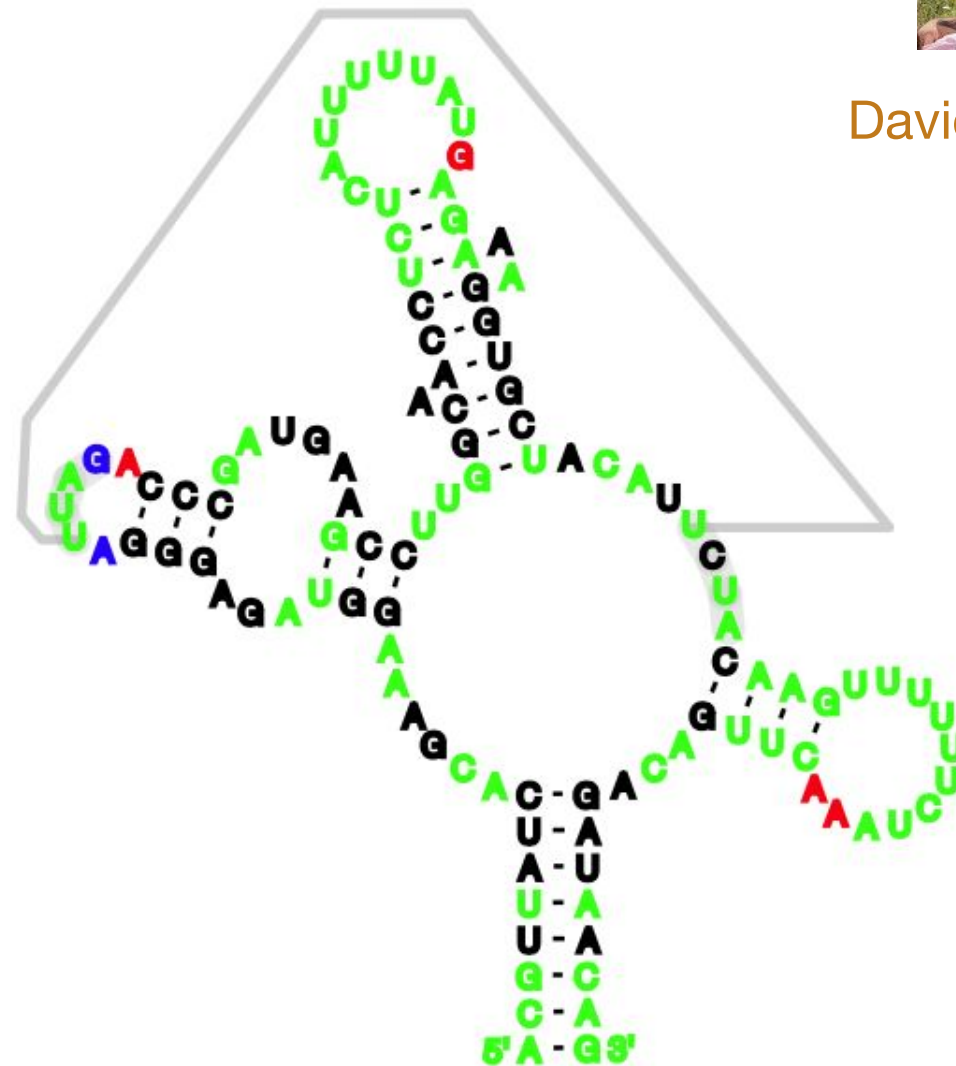
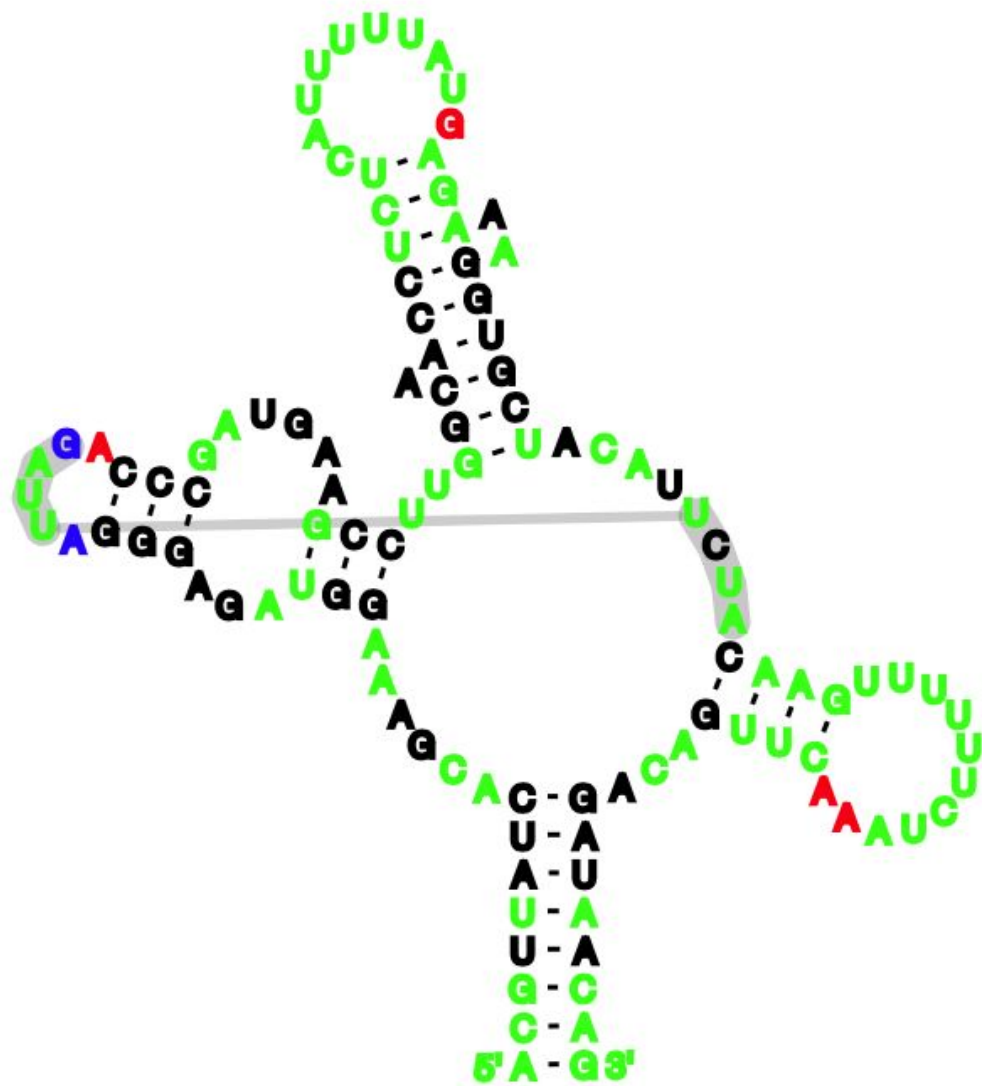
Use **RNArtist** and **RiboDraw** to author templates for R2DT and import using RNA 2D JSON Schema



Display pseudoknots using R2DT



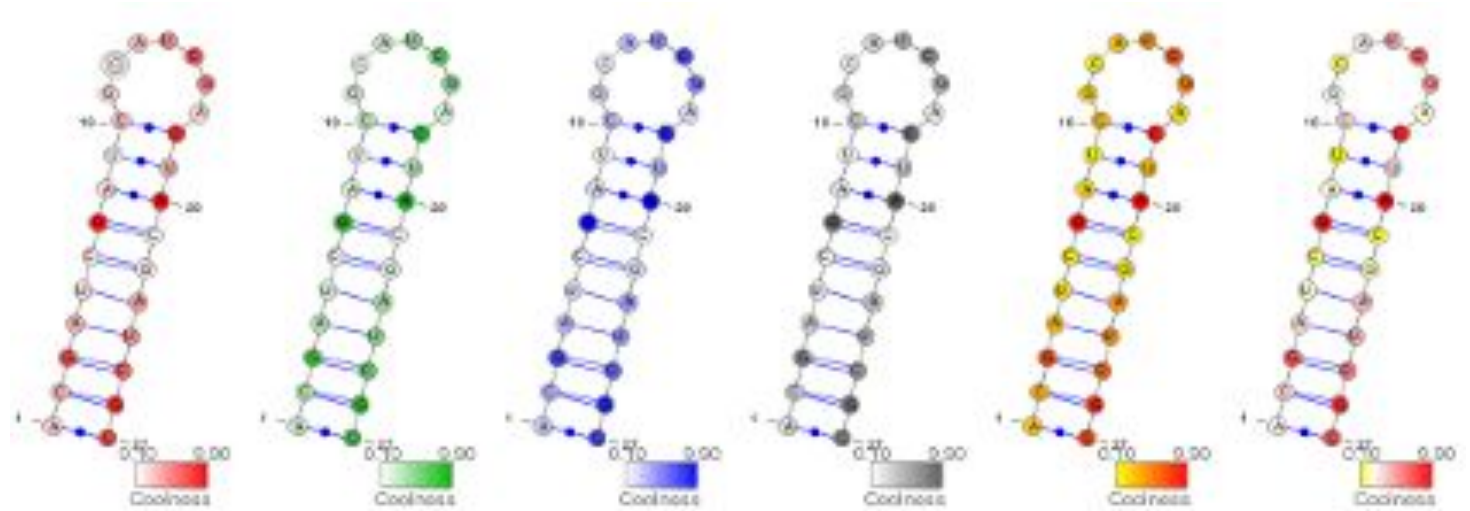
David Hoksza





Visualise data layers on top of secondary structures

- SHAPE reactivities
- SNPs
- Posterior probabilities
- Anything else!



<http://varna.lri.fr/index.php?lang=en&page=tutorial&css=varna>

Inspired by VARNA

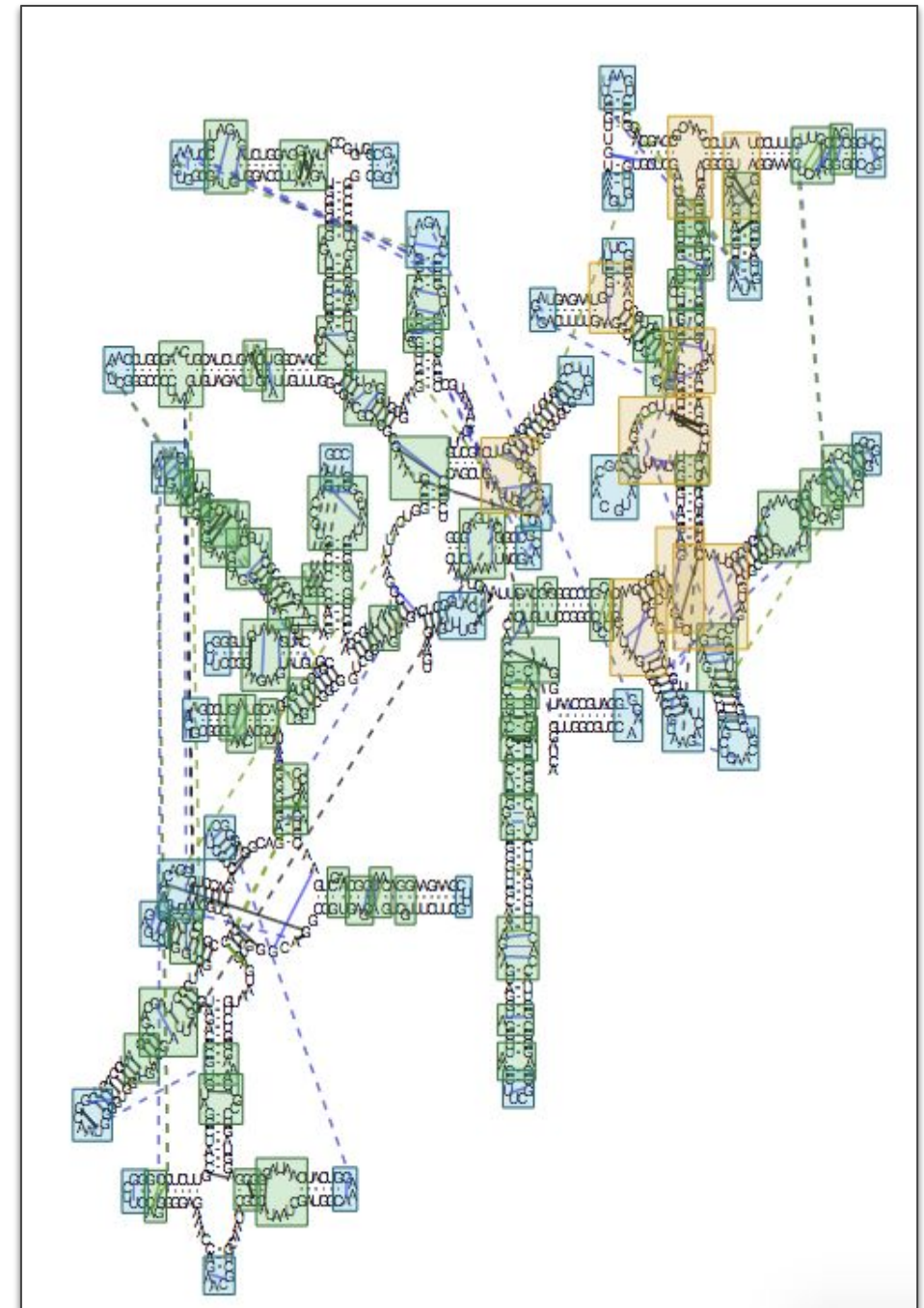


Display RNA 3D modules

- Inspired by the interactive airport diagrams from RNA 3D Hub
- Use JAR3D, BayesPairing, RMDetect, or any other method to identify modules and highlight them in 2D



Blake Sweeney





We have lots of work to do

- Interactive editor
- Pseudoknots
- Modules
- Data layers

Get in touch with your
wishlist!

Send us your **templates**
and **feedback!**

github.com/rnacentral/r2dt/issues
rnacentral.org/contact-us

contact@antonpetrov.com



HERE

THERE

EVERYWHERE

Sensible RNA structures...
everywhere!

- RNAcentral
- Rfam
- PDBe
- FlyBase
- GtRNAdb

and in the new papers!

The R2DT team

UCSC

Patricia Chan
Todd Lowe

BGSU

Craig Zirbel

NCBI

Eric Nawrocki

Independent

Anton I. Petrov

EMBL-EBI

Carlos Ribas
Fabio Madeira
Mandar Deshpande
Sri Devan Appasamy

Charles University

David Hoksza

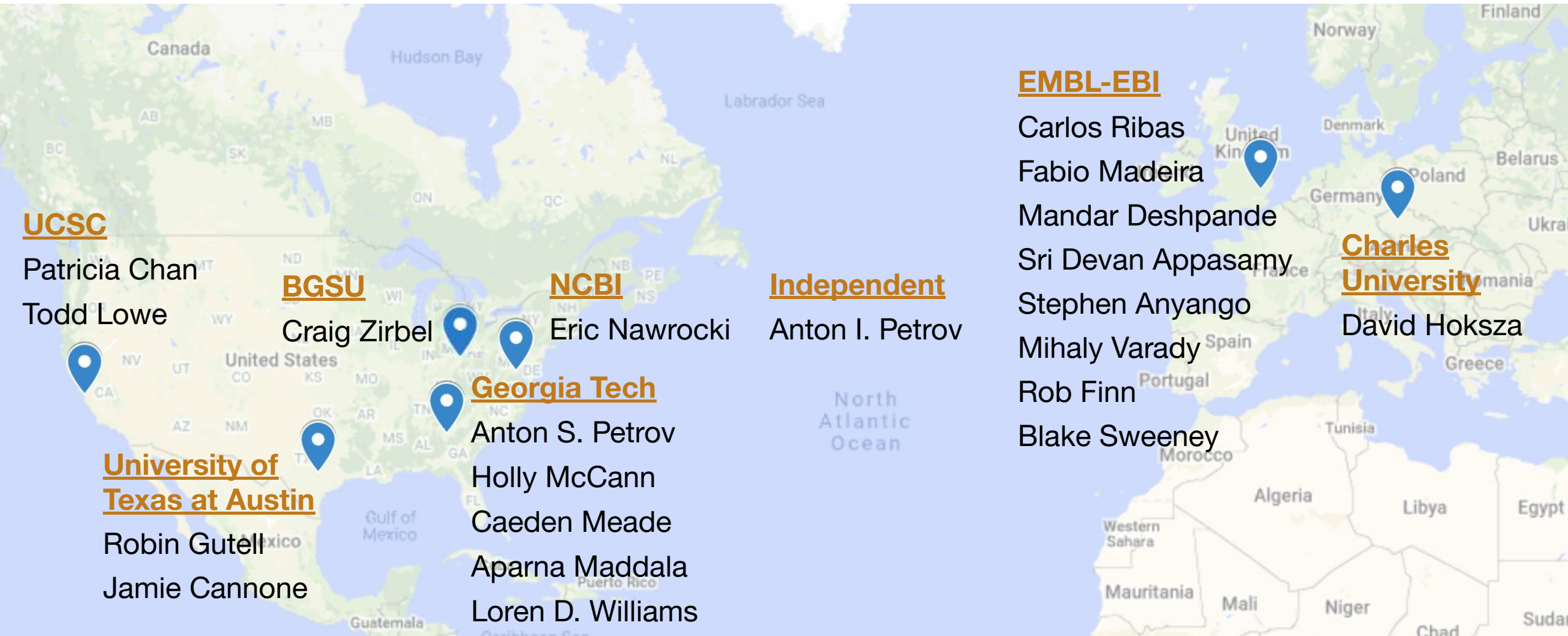
University of Texas at Austin

Robin Gutell
Jamie Cannone

Georgia Tech

Anton S. Petrov
Holly McCann
Caeden Meade
Aparna Maddala
Loren D. Williams

Rob Finn
Blake Sweeney



My Resume

• I ❤️ RNA

• Open to work or
collaboration

• Find out more at

• antonpetrov.com

•

•

Thank you!

**Please send your
templates
and feedback!**

github.com/rnacentral/r2dt/issues

rnacentral.org/contact-us

contact@antonpetrov.com

