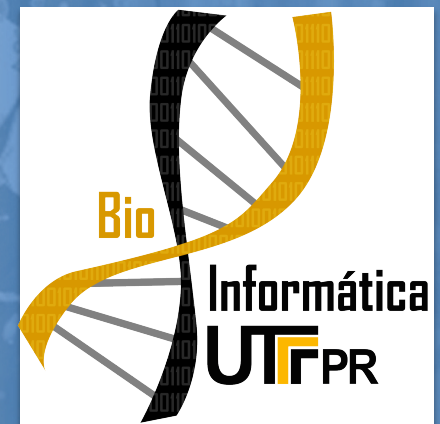


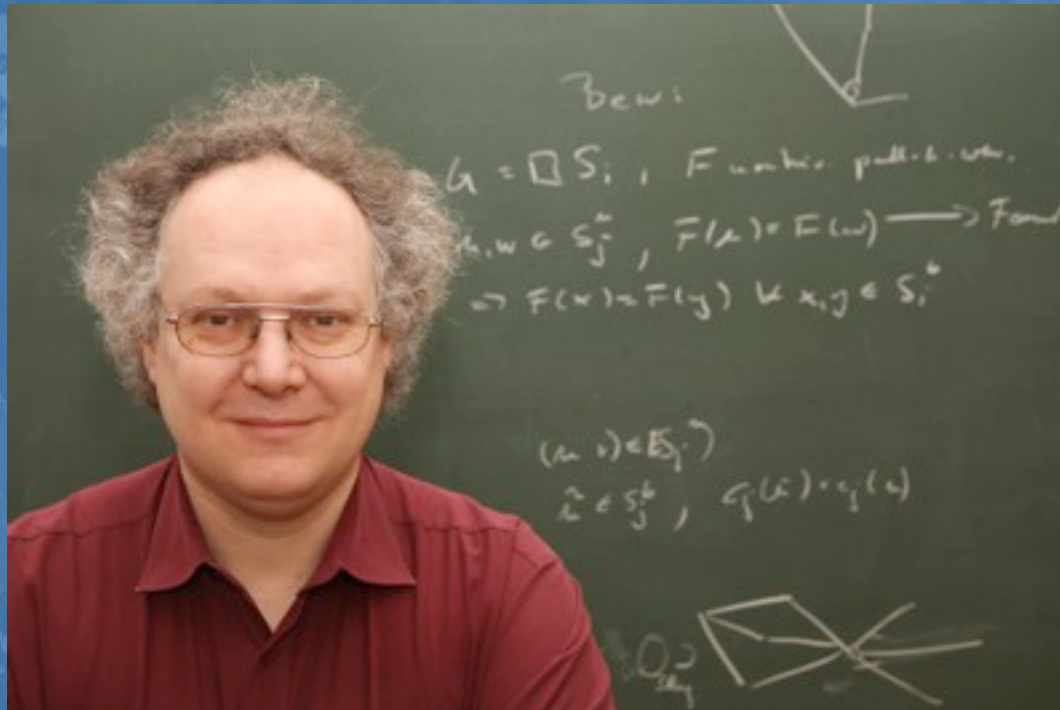
EXPLORE NON-CODING RNAS DATABASES KNOWLEDGE

Alexandre Rossi Paschoal
paschoal@utfpr.edu.br



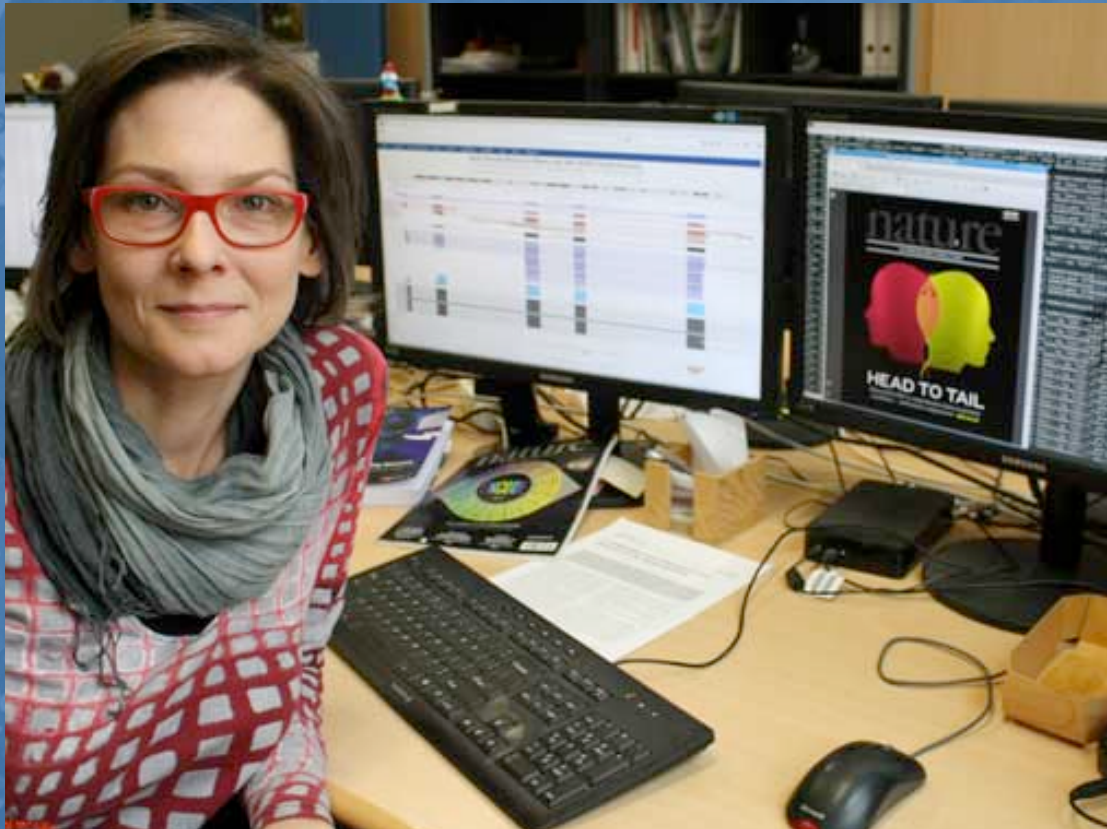
Before I start ...

- Thank you: Prof. Peter F. Stadler



Congratulation for the meeting

- Bled Organizers



What this talk is about this talk?



What this talk is about this talk?

- ncRNA information



What this talk is about this talk?

- ncRNA information:
 - where are they?
 - where is it available?
 - what is it possible to do?



In 2008 – PhD - Brazil

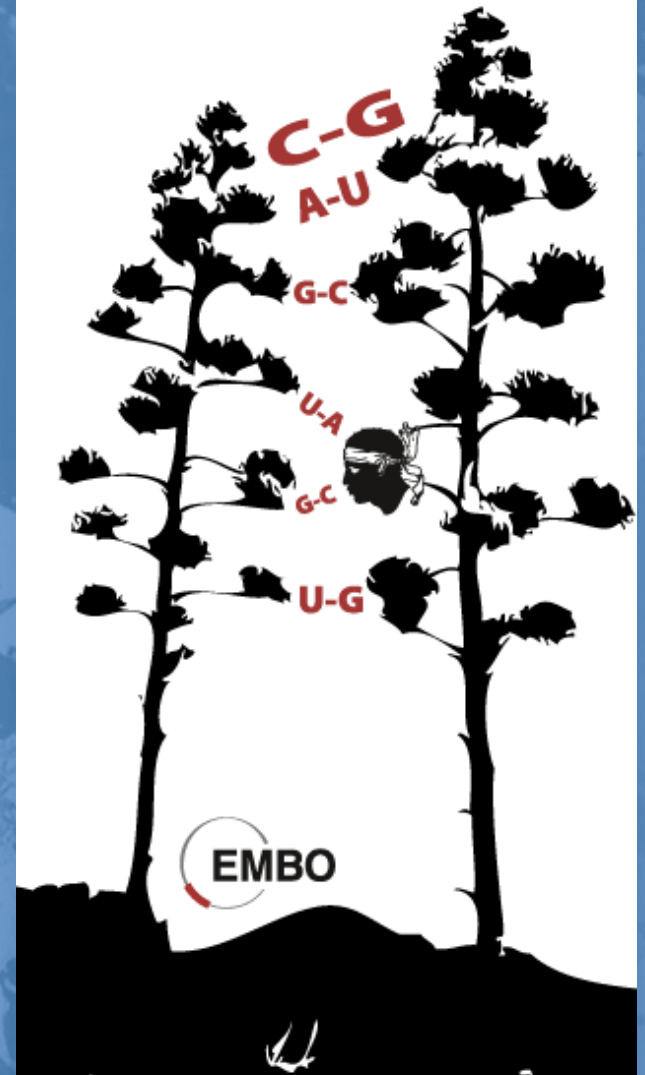
In 2008 – PhD - Brazil

- Kay Nieselt
- Ivo Hofacker
- Christoph Flamm
- Peter F. Stadler
- Robert Giegerich

EMBO Practical Course on Computational RNA Biology

April 26 – May 1, 2010

Institut d'Études Scientifiques de Cargèse, Corsica





ROMERO
BRITTO

An unique ncRNA repository?!

An unique ncRNA repository?!

- OK ... Databases: Rfam, miRBase and ... ??

An unique ncRNA repository?!

- OK ... Databases: Rfam, miRBase and ... ??
- But ...
- An unique repository of ncRNA data – WHERE?

- **2009** – Start this ideia
- **2012** – I finished – NRDR version 1.0



<http://www.ncrnadatabases.org/>

Other



Lapatas et al. *Journal of Biological Research-Thessaloniki* (2015) 22:9
DOI 10.1186/s40709-015-0032-5



Journal of
Biological Research

REVIEW

Open Access

Data integration in biological research: an overview



Vasileios Lapatas¹, Michalis Stefanidakis¹, Rafael C. Jimenez², Allegra Via³ and Maria Victoria Schneider^{4*}

Some Repositories



2012



2014 (2011)





A workflow for omic data analysis

10836 tools classified by omic technologies, applications and analytical steps

OMIC DATA

	Genomics		Epigenomics		Transcriptomics
	Proteomics		Metabolomics		Fluxomics
	Biochemical networks		Genotype-phenotype interactions		Genome editing
	RNA interference		Phenomics		Drug discovery
	Synthetic biology		Metabolic engineering		Biomolecular structure
	Health sciences		Immunology		Text-mining
	Miscellaneous		Bioinformatics infrastructure		

OMIC TECHNOLOGIES



High-throughput sequencing

Here, we surveyed bioinformatics software tools for high-throughput sequencing data analysis.



Mass spectrometry

Here, we surveyed bioinformatics software tools for the analysis of mass spectrometry data.



PCR

Here, we surveyed bioinformatics software tools for the analysis of polymerase chain reaction (PCR) data.



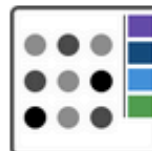
Flow cytometry & mass cytometry

Here, we surveyed bioinformatics software tools for the analysis of flow and mass cytometry data.



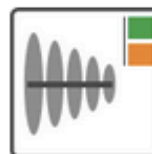
DNA fingerprinting

Here, we surveyed bioinformatics software tools for the analysis of DNA fingerprint images.



Microarray

Here, we surveyed bioinformatics software tools for the analysis of microarray data.



NMR spectroscopy

Here, we surveyed bioinformatics software tools for the analysis of nuclear magnetic resonance (NMR)...



Sanger sequencing

Here, we surveyed bioinformatics software tools for the analysis of Sanger sequencing data.



Bioimaging

Here, we surveyed bioinformatics software tools for the analysis of biological images.



Other omic technologies

Here, we surveyed bioinformatics software tools for the analysis of optical mapping and nCounter Analysis...



The non-coding RNA sequence database

Currently the RNAcentral Consortium is formed by **38** Expert Databases

Databases, **22** of which have already been imported into RNAcentral.

The RNAcentral logo consists of a stylized 'R' made of four colored squares (orange, purple, green, yellow) arranged in a 2x2 grid, followed by the text 'RNAcentral' in a blue, sans-serif font. The 'R' is partially overlapping the 'N'.

Text search

Search by *gene*, *species*, *publication*, *author* or any other keyword

[Browse sequences](#)

Sequence search

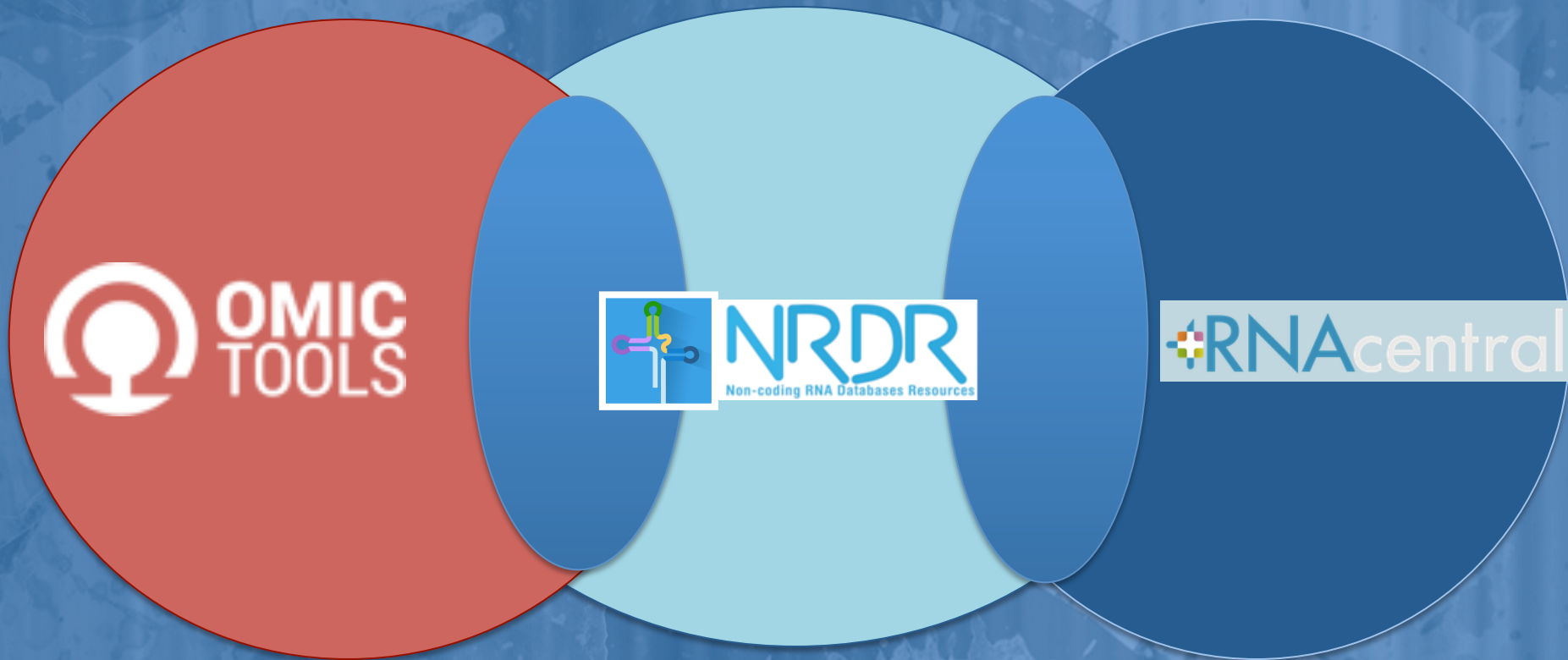
Search for similar sequences or look up your sequence in RNAcentral

[Search by sequence](#)

Genome browser

Explore RNAcentral sequences in your favorite genome locations

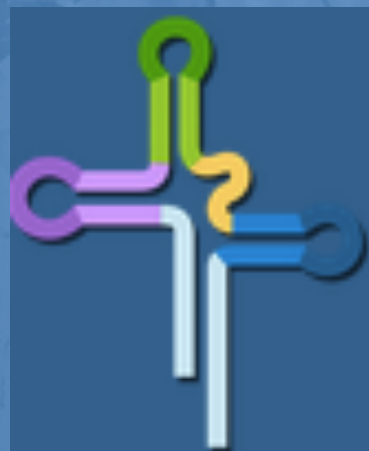
[Browse genomes](#)



**Omics
Catalog**

**Mix between
Both DBs**

**Integration with
mains (model)
ncRNAs databases**



NRDDR

Non-coding RNA Databases Resource

<http://www.ncrnadatabases.org/>

RNA Biology

Volume 9, Issue 3, 2012



Review

Non-coding transcription characterization and annotation

A guide and web resource for non-coding RNA databases

DOI: 10.4161/rna.19352

Alexandre Rossi Paschoal^{abc}, Vinicius Maracaja-Coutinho^{dbc},
 João Carlos Setubal^d, Zilá Luz Paulino Simões^e, Sergio Verjovski-
 Almeida^d & Alan Mitchell Durham^{f*}
 pages 274-282

Publishing models and article dates explained

Published online: 01 Mar 2012



Full text HTML



PDF



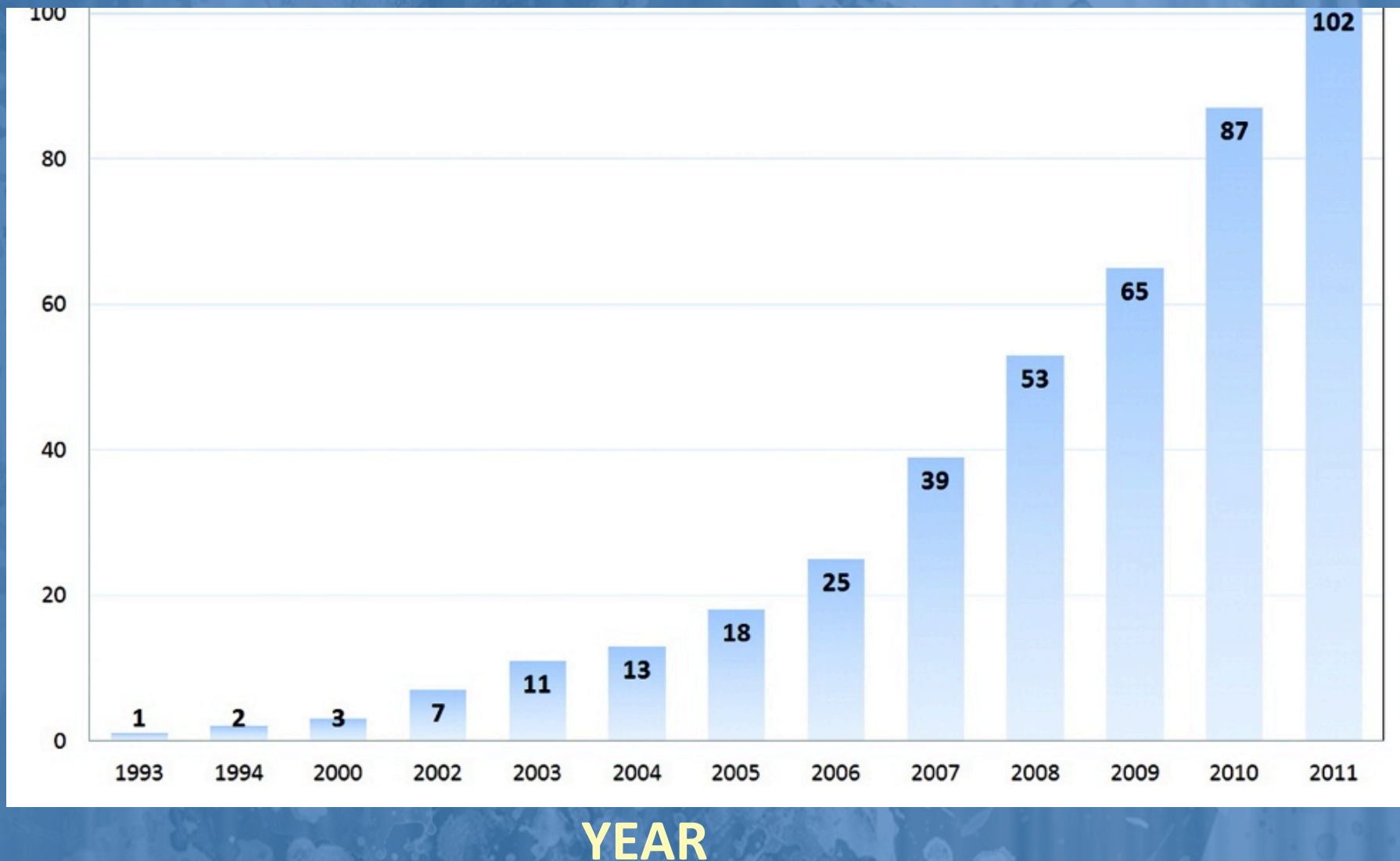
Supplemental

Free access

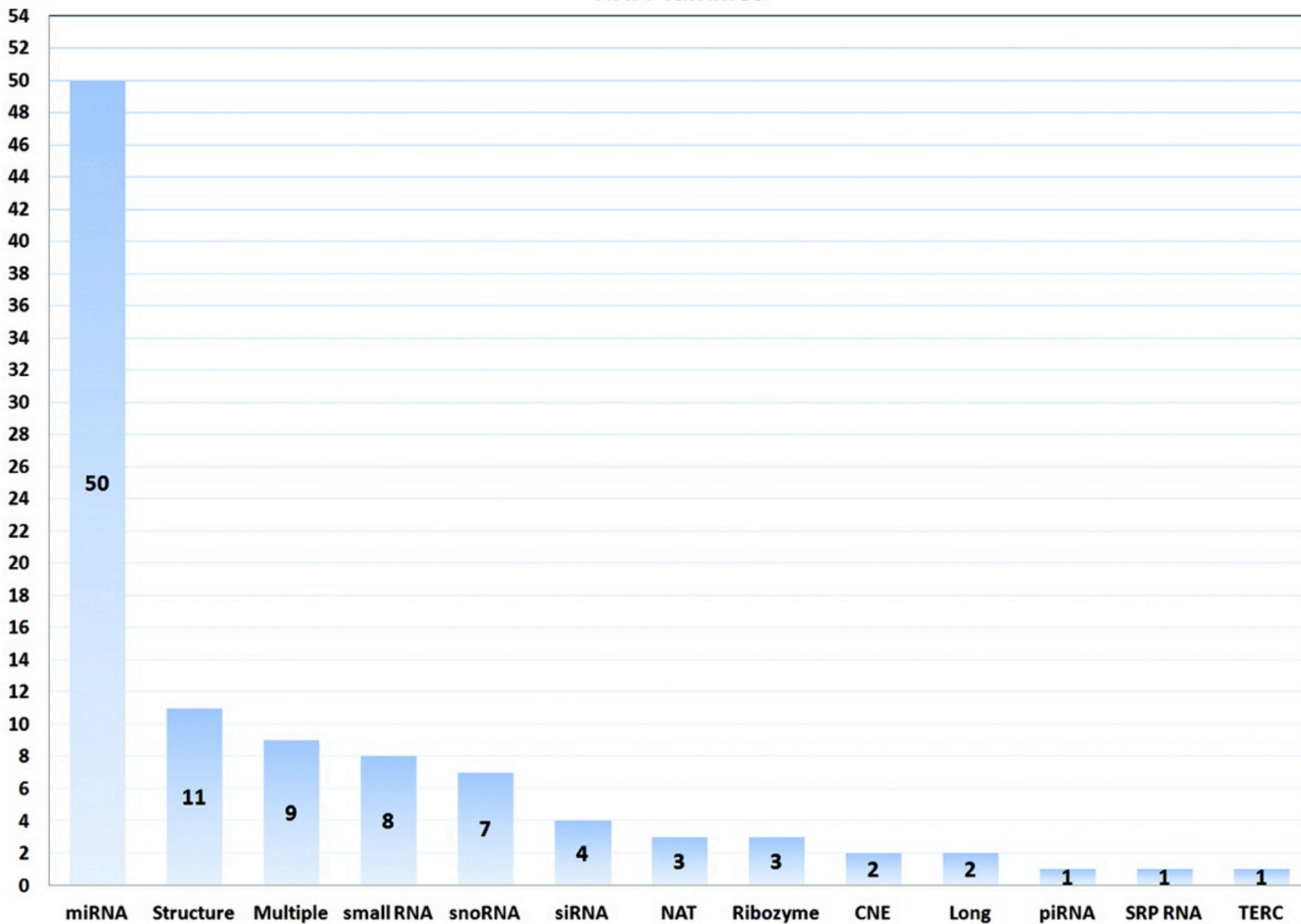
NRDR

- Version 1 – 2012 – 102 databases

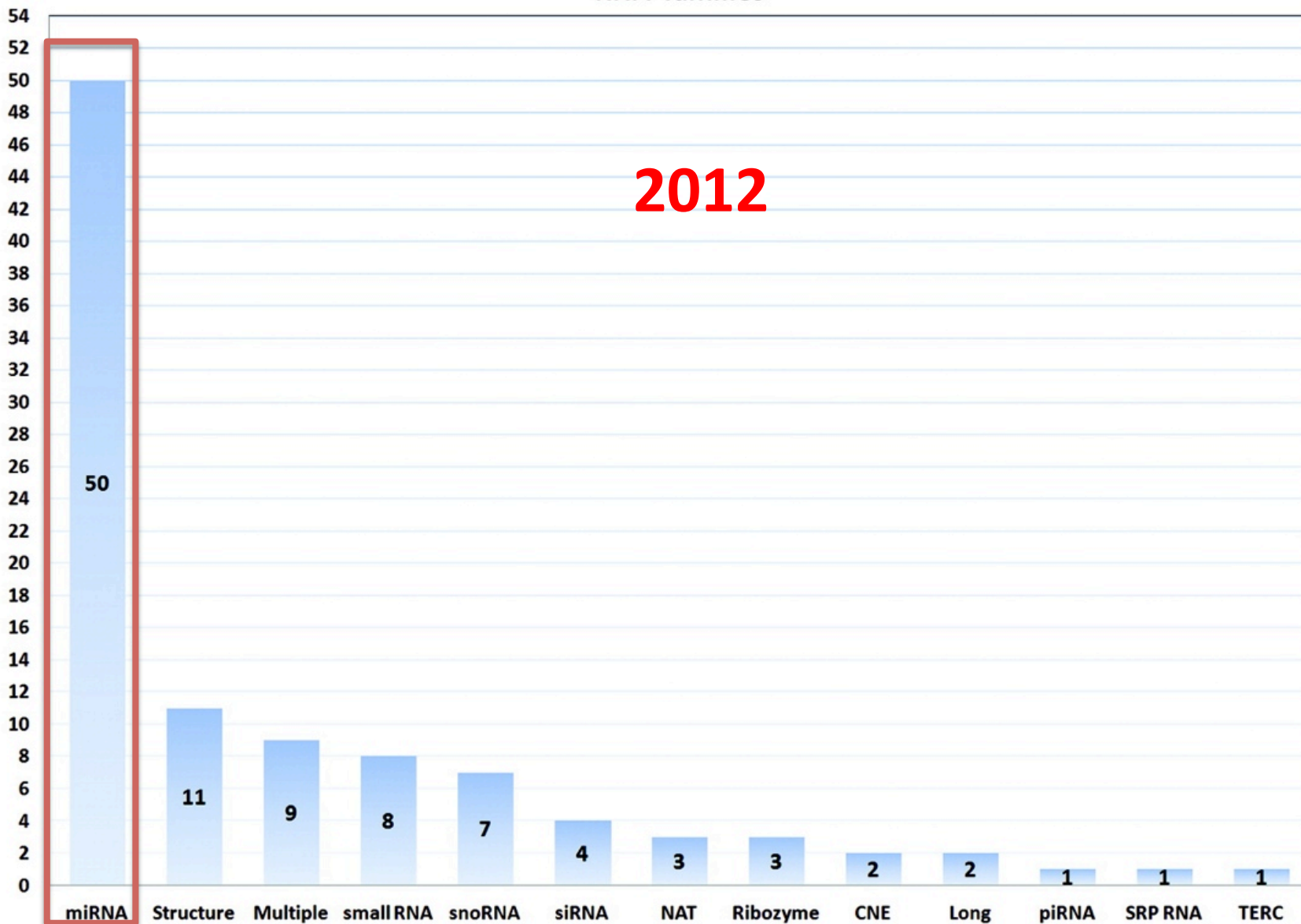
Cumulative Database Publication - Year



RNA families



RNA families



NRDR 2.0 – Some Numbers

- ~50% - miRNA
 - The majority using: *miRBase* data
- Specific: Longs RNAs DBs
 - 2012 – just 2 DBs
 - ***2016 – more than 10 databases (updating)***

NRDR

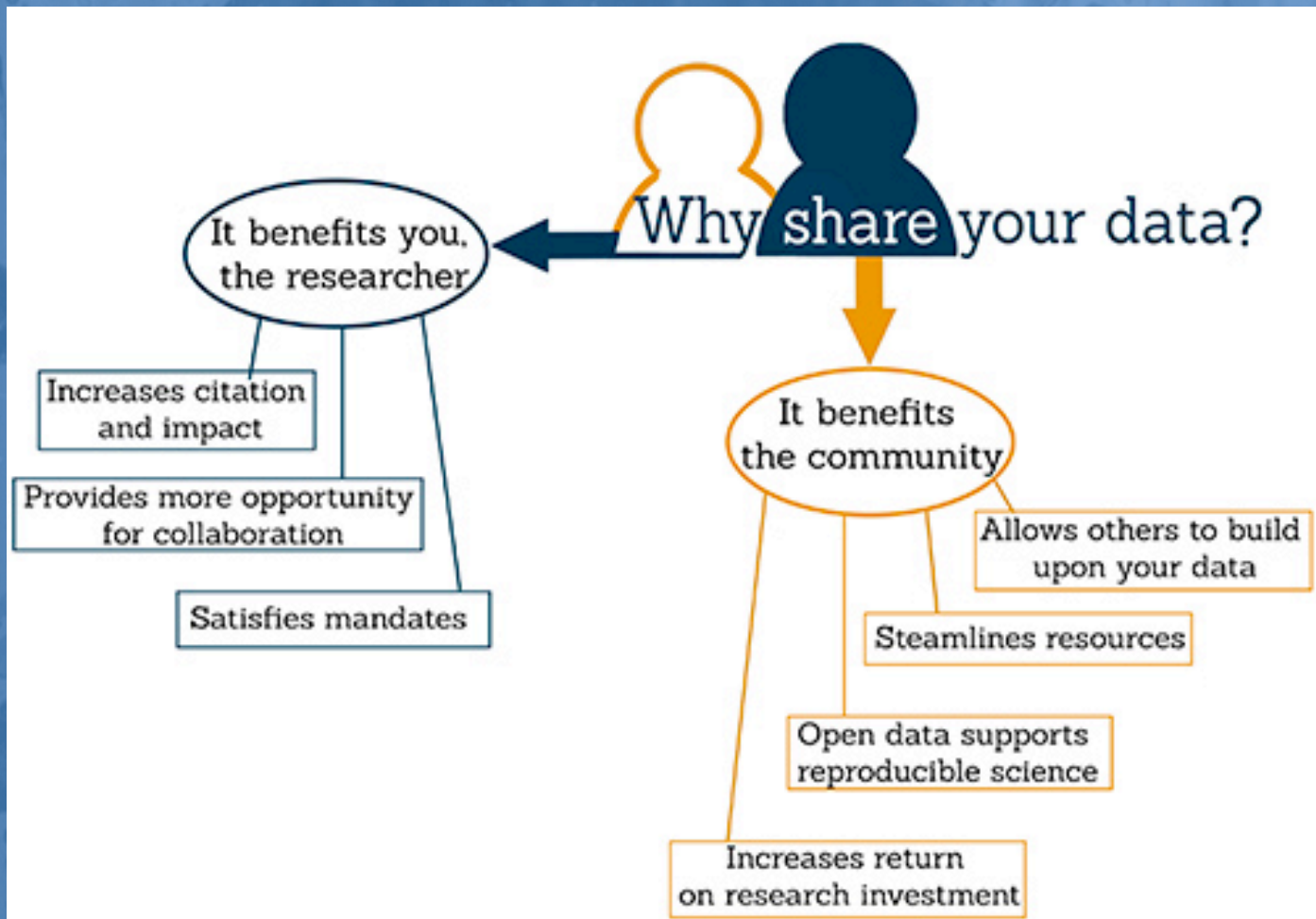
- 2015 - NRDR
 - Including 30 sequence databases available
 - * RNACentral: more option to explore the sequence
 - But not too much
- 2016 – Version 2 - 146 DB (working on)

“Funny” aspect

- NRDR – 140 DBs (2015 version)
 - **57%** make available any information (e.g. BED, GFF)
 - **Just ~28% FASTA**



Message 1: Share your data, Please



WHAT'S NEXT?



Future (Challenge) - NRDR

- Keep update what we have in NRDR
- OR/AND Describes new databases



Future (Problem) - DBs

- Not all databases keep updated



Perspectives – My interesting

- Explore this public data available
 - Not only describe the information or put sequence
 - Novel options/mechanisms
 - Novel visualization
 - What are these data telling me?!


Original Article

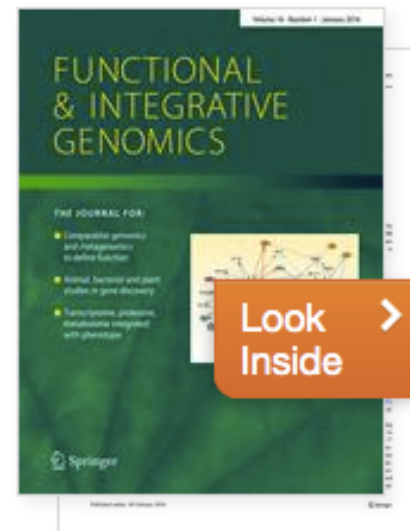
Functional & Integrative Genomics

pp 1-8

First online: 18 February 2016

PlanTE-MIR DB: a database for transposable element-related microRNAs in plant genomes

Alan P. R. Lorenzetti, Gabriel Y. A. de Antonio, Alexandre R. Paschoal, Douglas S. Domingues 



Message 2: suggestions or help us

- Suggestions - I will be very grateful
 - paschoal@utfpr.edu.br
- Help Us

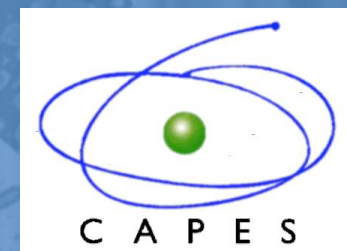
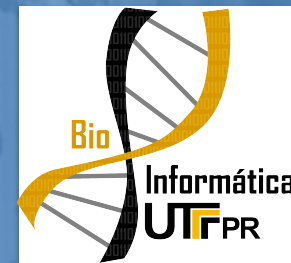


Final Message: Hard work



Vinicius Maracaja-Coutinho
João Carlos Setubal
Zilá Luz Paulino Simões
Sergio Verjovski-Almeida
Alan Mitchell Durham

Pedro Vinícius Borges Basseto
Liliane Santana Oliveira
Flávia Cal Sabino
Sibele Pinheiro de Souza
Douglas Silva Domingues
Artur Trancoso Lopo de Queiroz
Felipe Guimarães Torres
Raúl Arias-Carrasco



Grant: MCTI/CNPQ/Universal 14/2014 - Faixa A
Process: 454505/2014-0

Alexandre Rossi Paschoal

NRDR - www.ncrnadatabases.org

<http://bioinfo.cp.utfpr.edu.br/>
paschoal@utfpr.edu.br

Thank you for the attention





UNIVERSIDADE
TECNOLÓGICA FEDERAL
DO PARANÁ



23/09
104 anos

“De Escola de Aprendizes e Artífices (1909) à UTFPR (2013)”