



1920 | 2020

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO



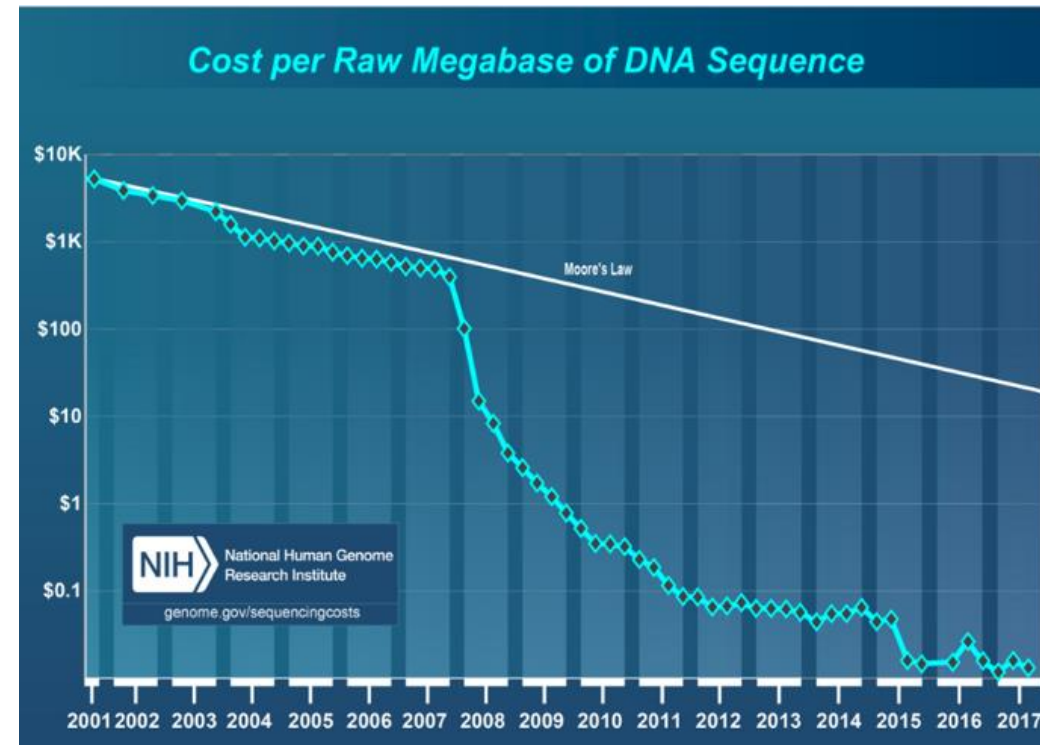
# LOW-BUDGET SCIENCE AND THE PRODUCTION OF COMPLETE AND NEARLY- COMPLETE MITOGENOMES OF *TEMNOTHORAX* ANTS USING PUBLIC DATA

Gabriel Alves Vieira & Francisco Prosdocimi

Laboratório de Genômica e Biodiversidade  
Instituto de Bioquímica Médica Leopoldo de Meis  
Universidade Federal do Rio de Janeiro

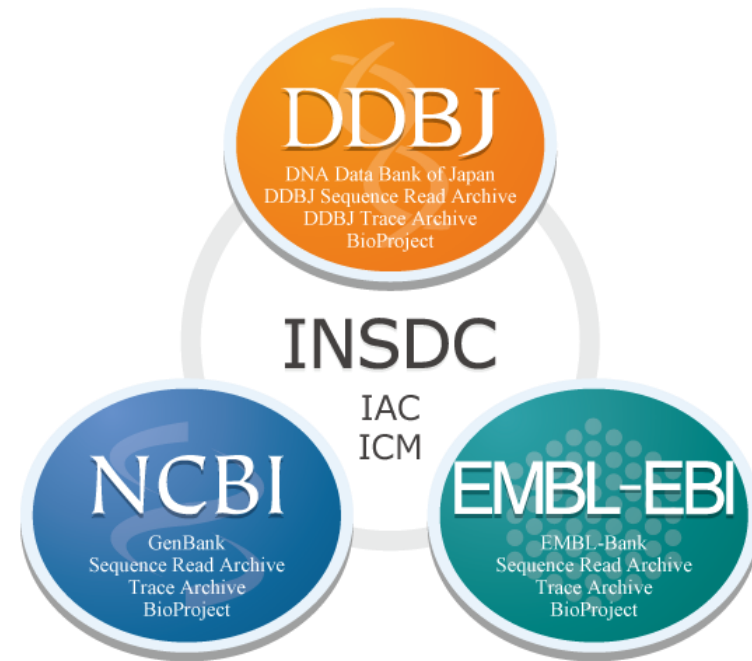
# Introduction - NGS and public data

- Next-Generation Sequencing (NGS)
  - ↑ genomic data produced
  - ↓ cost per sequenced base
- Journals: Incentivates raw data publication
  - Reproducibility
- Growth of public sequencing databases



# Sequence Read Archive (SRA)

- Part of an international collaboration (INSDC)
  - 3 databases:
    - SRA: NCBI Sequence Read Archive
    - ERA: EBI Sequence Read Archive
    - DRA: DDBJ Sequence Read Archive
  - Synchronized
- Several types of sequencing data:
  - DNA-seq; Bisulfite-seq; RNA-seq...
  - Enormous potential for “recycling”
  - Scientific advances at low costs

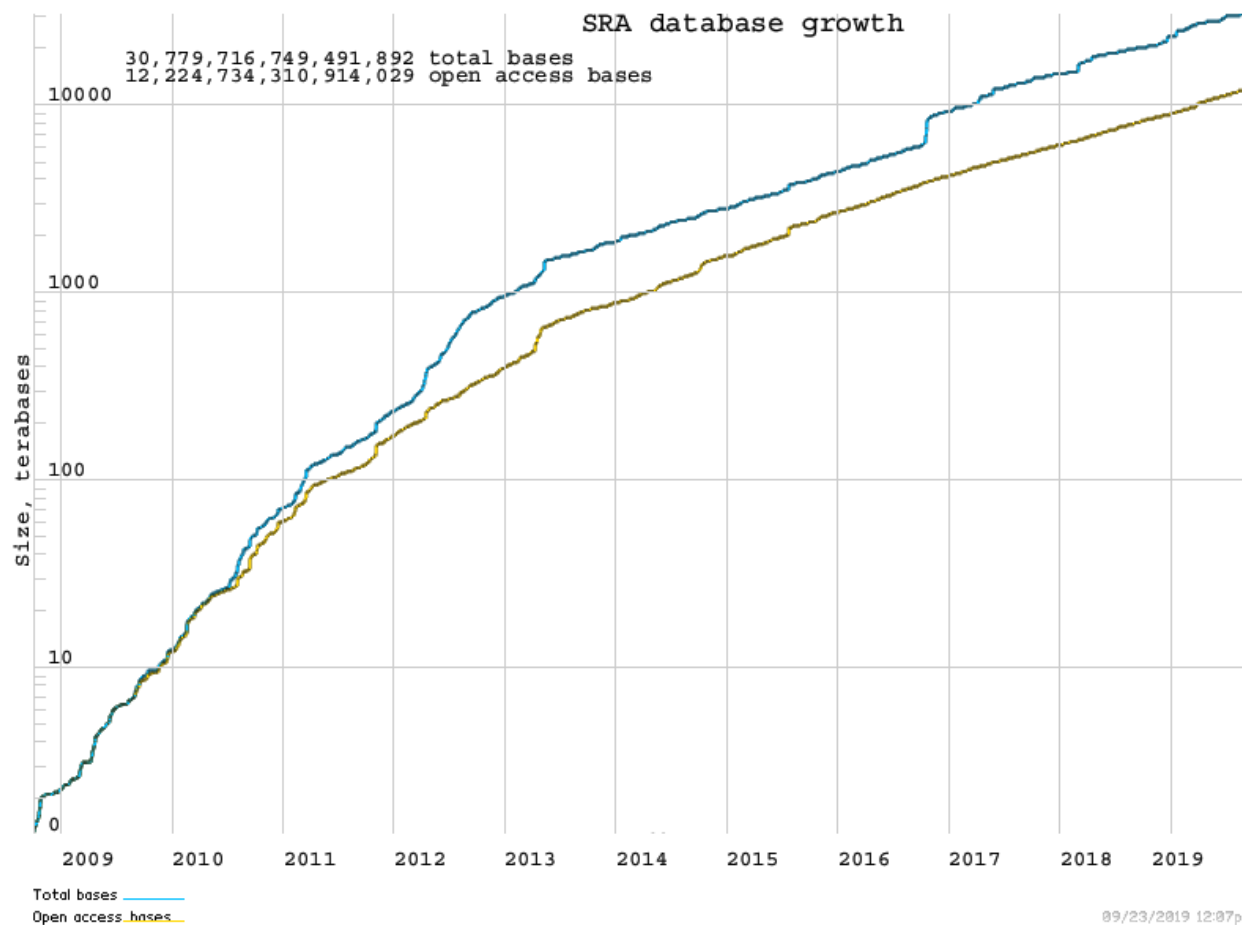


NEWS • 19 AUGUST 2019

# Brazil's budget cuts threaten more than 80,000 science scholarships

*If the country's main science-funding agency doesn't get more cash soon, young researchers will stop getting paid.*

Rodrigo de Oliveira Andrade

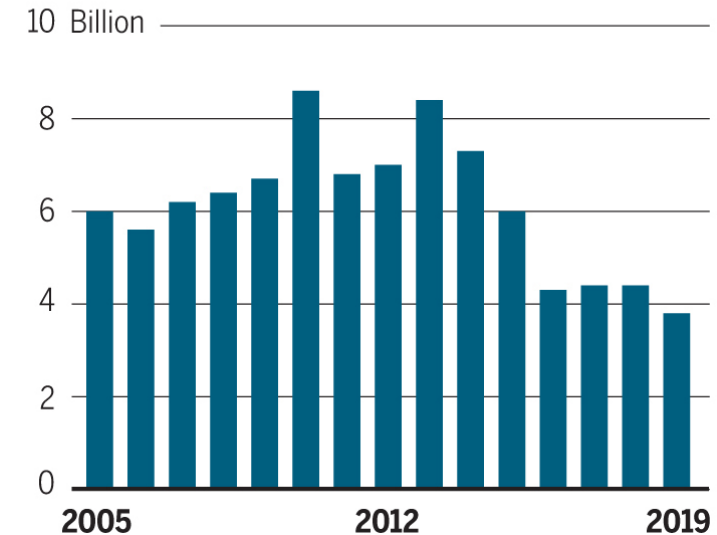


09/23/2019 12:07pm

## Science in decline

Even without the freeze, the Brazilian science ministry's budget for R&D—adjusted for inflation—has declined sharply the past few years.

## Budget in Brazilian reais



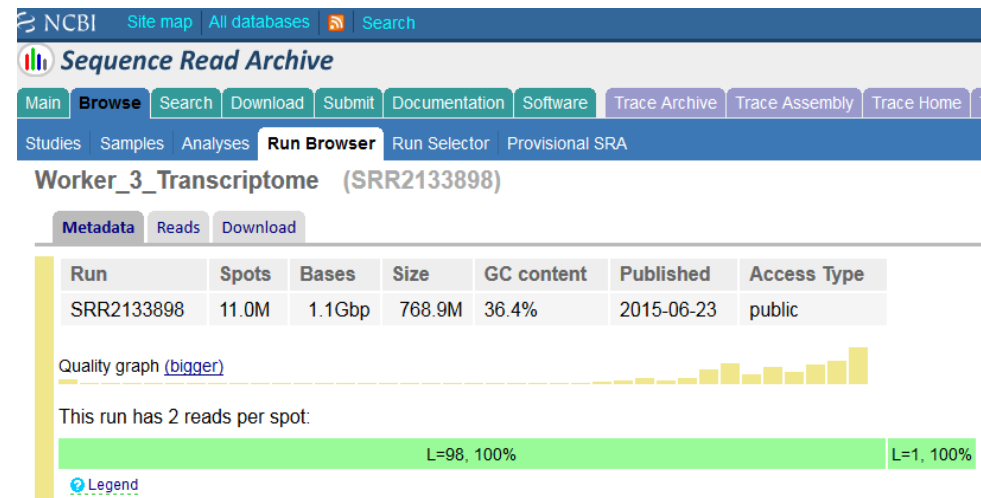
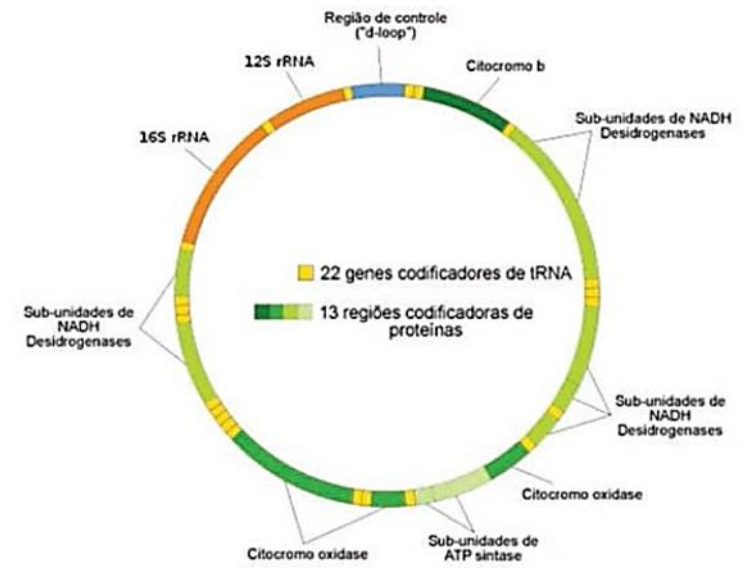
Brazil: ↓ Scientific funding

↑ Public sequencing data

How can these datasets foster new research?

# Public data and mitogenomics

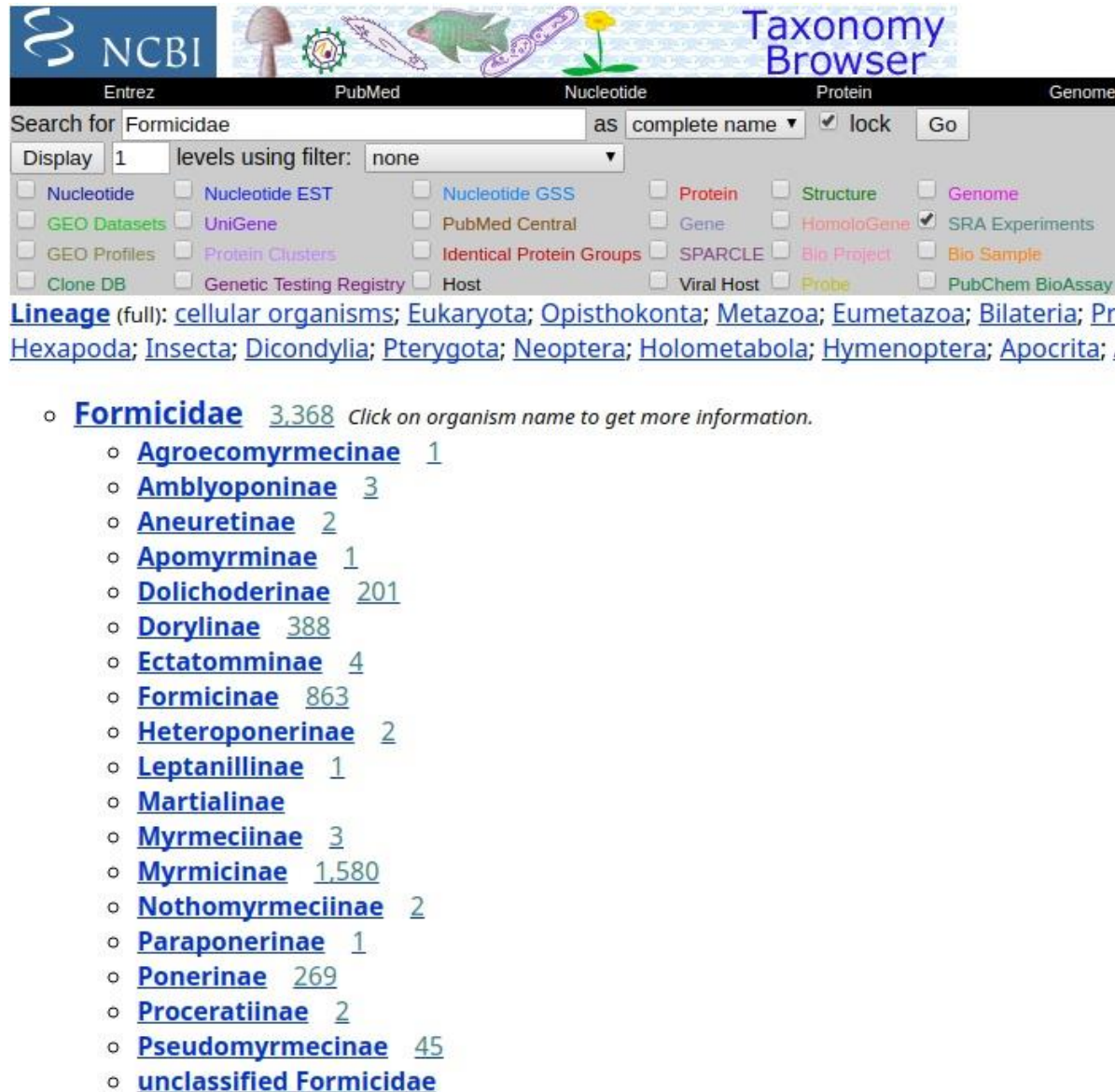
- Data used to answer a specific question
  - Not related to the mitogenome
- Sequencing
  - Mitochondrial sequences often present
  - 1 in each 200-1000 sequences is mitochondrial
  - Many species with public data available and no mitogenome published
- This data could be used to assemble the mitogenome
  - Phylogenetics (Krzemińska et al., 2017)
  - Phylogeography (Fields et al., 2018)
  - Populational genetics (Kilińcz et al., 2018)
  - Conservation (Rosel et al., 2017)





# Ants (Hymenoptera: Formicidae)

- More than 13000 *spp.*
- Huge amount of genomic sequencing data for the clade
  - $\approx 3.91$  Terabytes
- 35 complete mitogenomes
- Why not use this data?



The screenshot shows the NCBI Taxonomy Browser interface. The search bar contains "Formicidae" and the results are displayed as a list of taxa with their respective counts. The interface includes navigation tabs for Entrez, PubMed, Nucleotide, Protein, and Genome. The search results are organized into a hierarchical list, starting with the lineage and then listing the subfamilies of Formicidae.

NCBI Taxonomy Browser

Search for: Formicidae as complete name lock Go

Display: 1 levels using filter: none

Lineage (full): [cellular organisms](#); [Eukaryota](#); [Opisthokonta](#); [Metazoa](#); [Eumetazoa](#); [Bilateria](#); [Pr](#)  
[Hexapoda](#); [Insecta](#); [Dicondylia](#); [Pterygota](#); [Neoptera](#); [Holometabola](#); [Hymenoptera](#); [Apocrita](#);

- [Formicidae](#) 3,368 Click on organism name to get more information.
  - [Agroecomyrmecinae](#) 1
  - [Amblyoponinae](#) 3
  - [Aneuretinae](#) 2
  - [Apomyrminae](#) 1
  - [Dolichoderinae](#) 201
  - [Dorylinae](#) 388
  - [Ectatomminae](#) 4
  - [Formicinae](#) 863
  - [Heteroponerinae](#) 2
  - [Leptanillinae](#) 1
  - [Martialinae](#)
  - [Myrmeciinae](#) 3
  - [Myrmicinae](#) 1,580
  - [Nothomyrmeciinae](#) 2
  - [Paraponerinae](#) 1
  - [Ponerinae](#) 269
  - [Proceratiinae](#) 2
  - [Pseudomyrmecinae](#) 45
  - [unclassified Formicidae](#)

# Accessible molecular phylogenomics at no cost: obtaining 14 new mitogenomes for the ant subfamily Pseudomyrmecinae from public data

Gabriel A. Vieira and Francisco Prosdocimi

Instituto de Bioquímica Médica Leopoldo de Meis, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro, Brazil

Submitted 7 November 2018  
Accepted 10 December 2018  
Published 24 January 2019

Corresponding authors  
Gabriel A. Vieira



## Mitochondrial DNA Part B Resources

ISSN: (Print) 2380-2359 (Online) journal homepage: <https://www.tandfonline.com/loi/tmdn20>

## Complete mitochondrial genomes for three lizards (*Anolis punctatus*, *Sceloporus woodi*, and *S. grammicus*): a contribution to mitochondrial phylogenomics of Iguanoidea

Ana Teresa Nogueira Dumans, Gabriela Warwar Teixeira, Gabriel Alves Vieira, Deise Schroder Sarzi, Carolina Furtado, W. Bryan Jennings & Francisco Prosdocimi

# *Temnothorax* spp. – Myrmicinae subfamily

- Highly diverse - 408 spp. (AntCat)  
Unclear taxonomy and evolution
- Model organism – Social insects  
Ecology (Blatrix & Herbers, 2003)  
Behaviour (Charbonneau et al. 2015)
- **No mitogenome described**



Prebus *BMC Evolutionary Biology* (2017) 17:250  
DOI 10.1186/s12862-017-1095-8

BMC Evolutionary Biology

RESEARCH ARTICLE

Open Access



Insights into the evolution, biogeography  
and natural history of the acorn ants, genus  
*Temnothorax* Mayr (hymenoptera:  
Formicidae)

Matthew Prebus 

Prebus (2017)

UCE Datasets:

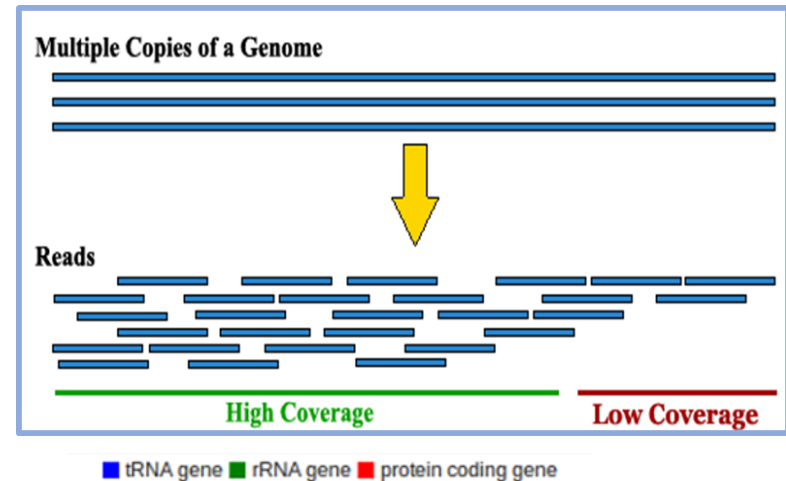
37 *Temnothorax* spp.  
13 outgroups

**How to use these  
datasets?**



# Material and Methods

- Downloaded data from ERA (SRA file format)
- Converted data to fastq
- Assembly
  - NOVOPlasty (primary assembly)
  - MITObim (close gaps and refine assembly)
- Assembly check: TABLET
  - Coverage
    - How many sequences corroborates the assembly
  - Circularization
- Anotation:
  - Identification of specific regions in a genome (Ex. Genes)
    - Automatic: MITOS Web Server
    - Manual: Artemis (NOT FINISHED)



**MITOS WebServer**

Name:

Email:

Job Identifier:

Genetic Code\*:

Fasta File\*:

\* = required

A tutorial on how to use MITOS, including an [example](#) and the used sample data, can be found [here](#).



# Phylogenomics

- Mitogenomes:
  - *Temnothorax* spp.
  - Ants previously described
  - Outgroups (bees)
- Supermatrix:
  - 13 protein-coding genes
  - Aligned and concatenated
- Phylogenomics inference:
  - Neighbor Joining
  - Resampling: Bootstrap (1000 replicates)



DNA alignments

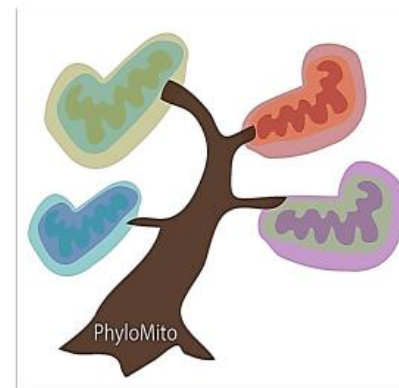
Tree inference  
(same study)

Concatenation



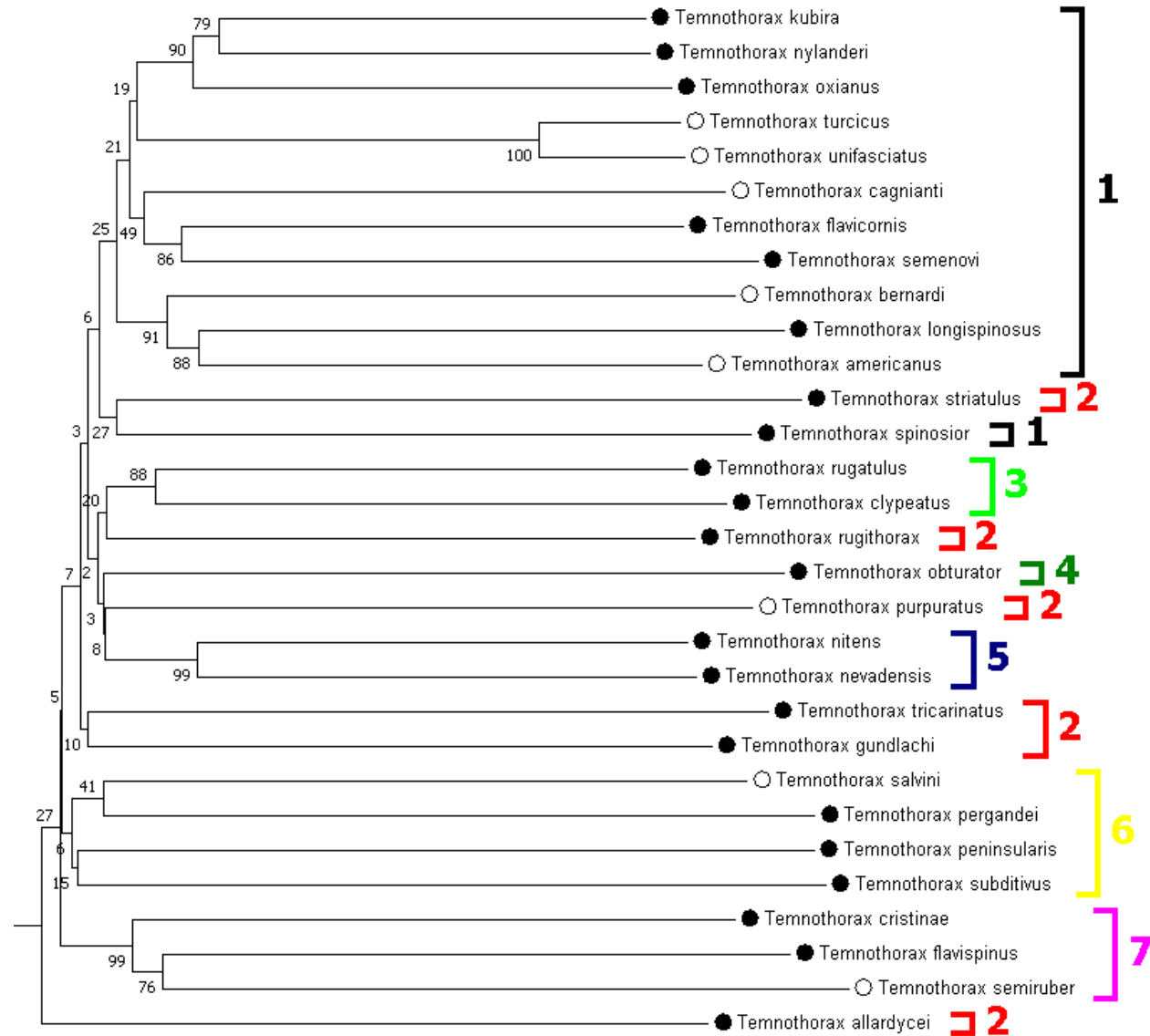
Supermatrix

Tree  
inference  
(e.g. ML, Bayesian)



# Results

- *Temnothorax* spp.:  
30 mitogenomes  
22 complete  
8 partial (nearly complete)
- Monophyletic genus
  - Highly supported
- Basal nodes
  - Low bootstrap support
  - Paraphyly
    - 2 out of 7 clades



## Mitogenomes:

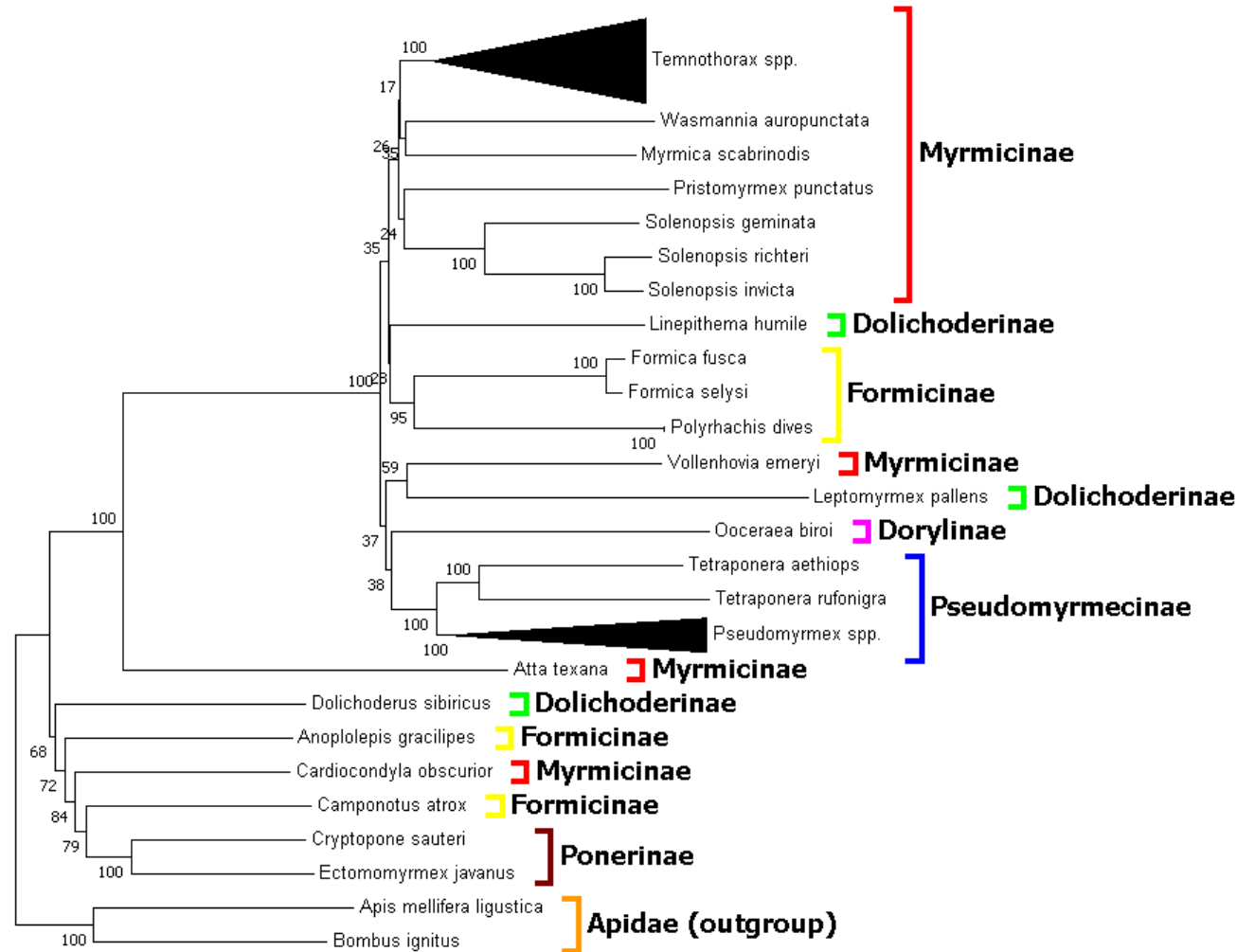
- Complete
- Partial

## Clades:

- |                 |                 |
|-----------------|-----------------|
| 1 - Palearctic* | 5 – andrei      |
| 2 – sallei*     | 6 – salvini     |
| 3 – rugatulus   | 7 – rottenbergi |
| 4 – obturator   |                 |

# Results

- **Formicidae Tree:**
- Monophyletic clades
  - Pseudomyrmecinae
  - Ponerinae
  - All genera
  - 100% BS support
- Low bootstrap values
  - Myrmicinae nodes





# Discussion

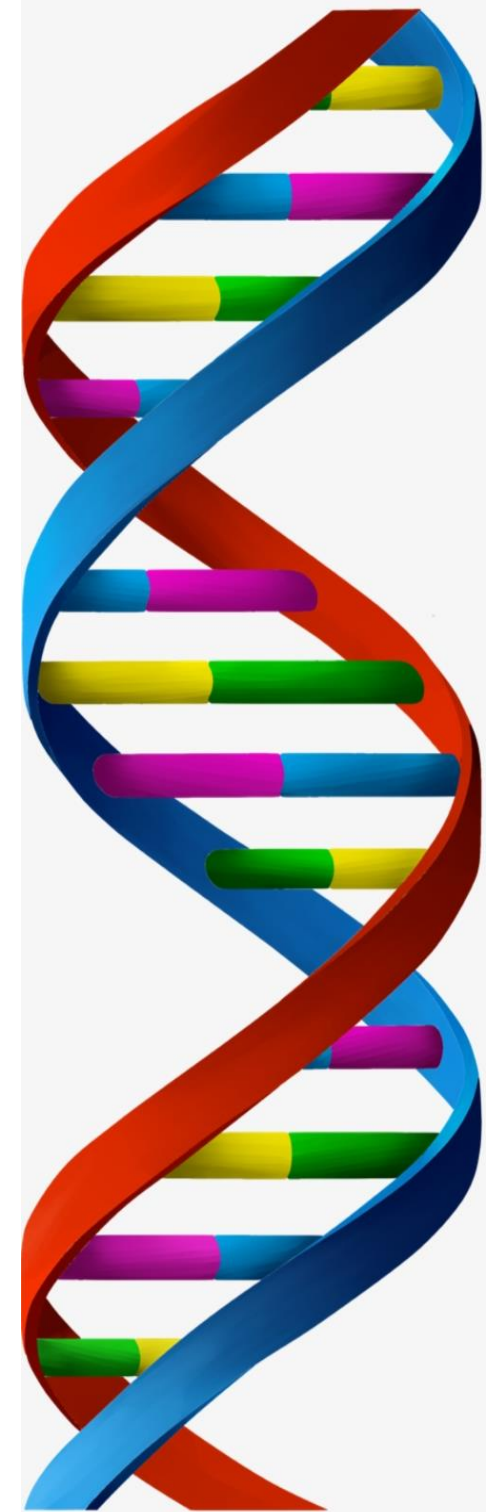
- Overall:
  - Moderate phylogenetic inference accuracy
  - All genera monophyletic
  - Good resolution – resolve shallow nodes
- Low bootstrap values
  - Mostly in Myrmicinae – hyperdiverse
  - Improved taxonomic coverage
  - More robust tree inference methods
  - Resolve paraphyletic relationships
  - Also applicable to poorly support nodes in *Temnothorax* genus
- Next steps:
  - Mitogenome assembly (outgroup species)
  - Manual curation
  - Maximum Likelihood
  - Synteny and comparative genomics analyses





# Conclusion

- This preliminary study:
  - Obtained and analyzed *Temnothorax* mitogenome sequences for the first time
  - No sequencing costs
    - Illumina 'partial genome' run:  $\approx$  2.000 R\$ per sample
    - 30 mitogenomes: **60.000 R\$ saved!**
  - Free bioinformatics software
  - 4 GB RAM – No supercomputer required
  - **NO/LOW BUDGET MITOGENOMICS**
  - MitoFree – Automatic assembly tool
- Contribution to ant mitogenomic resources
- Highlight the untapped potential of public databases
- Inspire similar works
  - ↑ mitogenomes : Evolution and Conservation (Non-model)



# REFERENCES

- Blatrix R, Herbers JM. Coevolution between slave-making ants and their hosts: host specificity and geographical variation. *Mol Ecol*. 2003;12:2809–16.
- Charbonneau D, Hillis N, Dornhaus A. 'Lazy' in nature: ant colony time budgets show high 'inactivity' in the field as well as in the lab. *Insect Soc*. 2015;62:31–5
- Fields, Peter D., et al. "Mitogenome phylogeographic analysis of a planktonic crustacean." *Molecular Phylogenetics and Evolution* (2018).
- Kılınç, Gülşah Merve, et al. "Investigating Holocene human population history in North Asia using ancient mitogenomes." *Scientific reports* 8.1 (2018): 8969.
- Krzemińska, Urszula, et al. "Population mitogenomics provides insights into evolutionary history, source of invasions and diversifying selection in the House Crow (*Corvus splendens*)." *Heredity* 120.4 (2018): 296.
- Rosel, Patricia E., et al. "Examining metrics and magnitudes of molecular genetic differentiation used to delimit cetacean subspecies based on mitochondrial DNA control region sequences." *Marine Mammal Science* 33.S1 (2017): 76-100.

# OBRIGADO!

Email: [gabriel.vieira@bioqmed.ufrj.br](mailto:gabriel.vieira@bioqmed.ufrj.br)

