

# **FILOGENÔMICA MITOCONDRIAL EM FORMIGAS (HYMENOPTERA: FORMICIDAE)**

**Aluno:** Gabriel Alves Vieira

**Orientador:** Francisco Prosdocimi



```
AACTCAAGCTGCTCTGGGGCTGTGATCTGCCCTCAAACCCACAGCCTGGGTAGCAGG  
AGGACCTTGATCTCCTGGCACAGATGAGGAATCTCTTCTCTGCTTGAG  
GACAGACATGACTTTGGATTCCCCAGGAGGAGTTGCCAACAGTTCCAAAAGGCT  
GAAACCATCCCTGTCCTCCATGAGATGATCCAGCAGATCTCAATCTCTCAGCACA  
AAGGACTCATCTGCTGCTGGGATGAGACCTCTAGACAAATTCTACACTGAACCT  
TACAGCAGCTGAATGACCTGGAAGCCTGTTGATACAGGGGTGGGGTAGCAGAG  
ACTCCCCTGATGAAGGAGGACTCCATTCTGGCTGTGAGGAATACTCTCAAAGAATC  
ACTCTATCTGAAAGAGAAGAAATACAGCCCTTGCCCTGGGAGGTGTCAGAGCA  
GAAATCATGAGATCTTTCTTGCAACAAACTGCAAGAAAGTTAAGAAGTAAG  
GAATGA, TGTGATCTGCCCTCAAACCCACAGCCTGGGTAGCAGGAGGACCTTGATGC  
TCTGGCACAGATGAGGAGAATCTCTTCTCTGCTTGAGGACAGCATGACT  
TTGGATTTCCCAGGAGGAGTTGGCAACCAGTCCAAAAGGCTGAAACCATCCCTG  
TCCCTCATGAGATGATCCAGCAGATCTCAATCTCTCAGCACAAAGGACTCATCTG  
CTGCTGGGATGAGACCTCTAGACAAATTCTACACTGAACCTACAGCAGCTGA  
ATGACCTGGAAGCCTGTTGATACAGGGGTGGGGTAGCAGAGACTCCCTGATGC
```

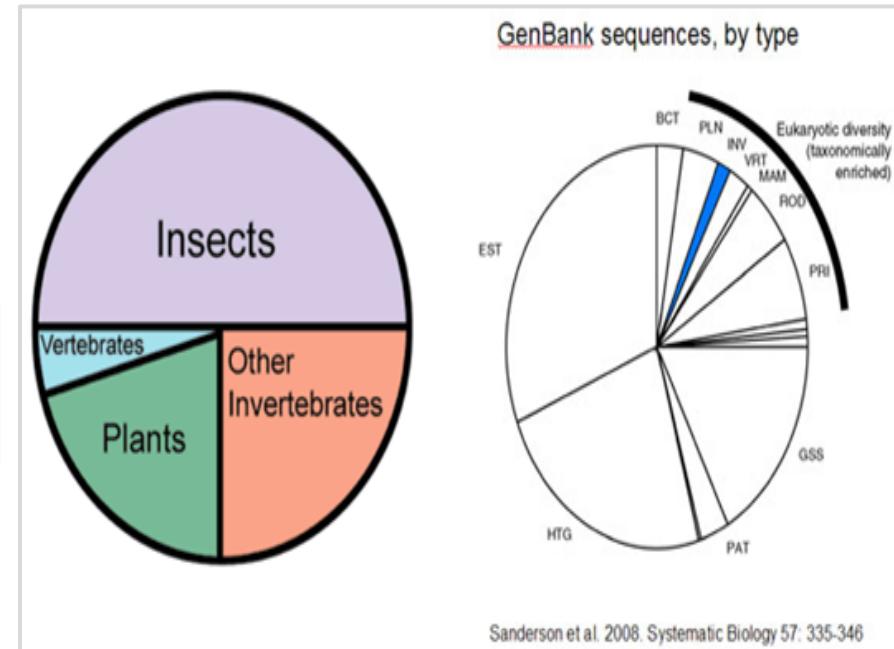
# Carga horária

- Bioética e Biossegurança – 30h
- Fundamentos em Macromoléculas – 45h
- Fundamentos em Biologia Molecular – 45h
- Fundamentos em Metabolismo – 45h
- Iniciação à Ciência: 2º Grau I – 90 h
- No-Budget Science – 15h
- Big Data - 30h
- Ciência e arte – 75h
- Estratégias de Divulgação Científica - 45h
- **24/25 ALVs – Seminários Plenos – 30h**
- **20/25 Journals - Seminários Gerais – 30h**
- **Tópicos Avançados – 15h**
- **Atividades didáticas I -45h**
- **TOTAL: 540h**



# Formigas (Hymenoptera: Formicidae)

- Engenheiras ecossistêmicas
  - Alteram características do solo
  - Serviços ecossistêmicos
- Importância econômica
  - Pragas
  - Agentes de controle biológico
- Grande biodiversidade
  - $\approx 13.000$  spp.
  - GenBank
    - 18 genomas parciais



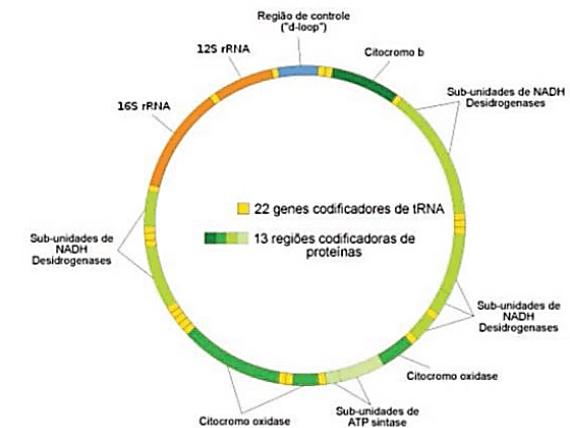
# Filogenômica mitocondrial

- Mitogenomas

- Molécula circular
- Sem recombinação
- Herança materna
- Amplamente usados em reconstruções filogenéticas
- 14 mitogenomas completos de formiga
  - Apenas 3 das 17 subfamílias existentes

- LAMPADA

- Excelência em filogenômica mitocondrial
- Elaboração de ferramentas específicas
  - ↓ tempo gasto - montagem e análise
- 30+ mitogenomas publicados
  - Aves, moluscos, insetos, dentre outros



## Mitochondrial Phylogenomics in Animal Models

Francisco Prosdocimi, Marcela Uliano-Silva, Igor Rodrigues da Costa, Nicholas Costa Barroso Lima



Automatic assembly and annotation of whole mitochondria

- >> DeNovo assembly using 5 different kmers
- >> Reference-based assembly
- >> Automatic annotation
- >> GenBank annotated output



Download all mitochondria from a taxID

mt-dowloader

- \$> Finds a taxID on NCBI
- \$> Searches for whole mitochondria on RefSeq
- \$> Download all mitochondria in GB format



mitoCompare

Mitochondria comparison and reports

- >> Multiple sequence alignment
- >> Definition of a query sequence
- >> Amino acid mutations tabulated



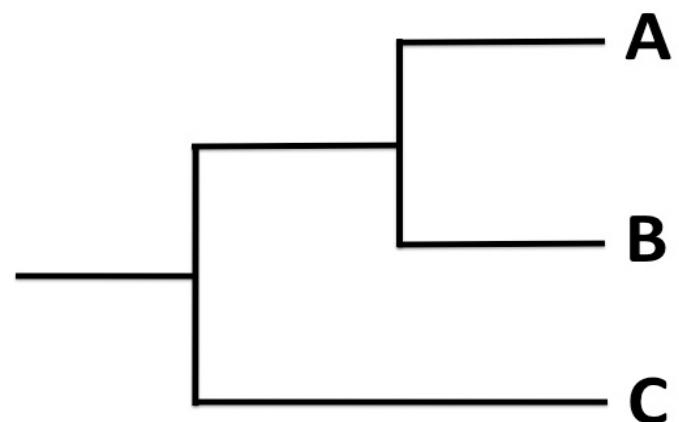
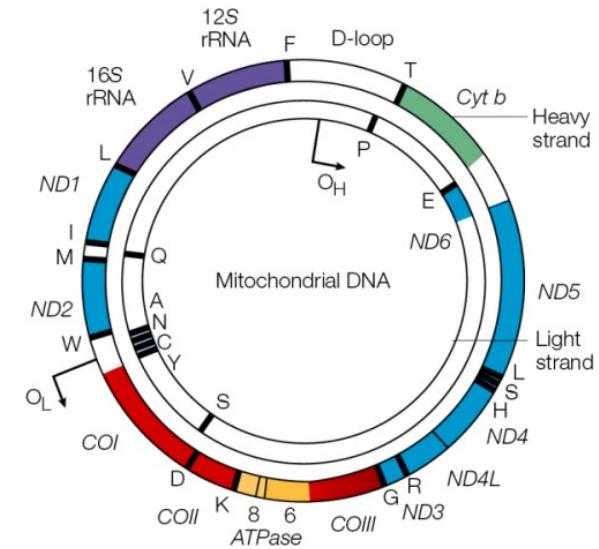
phyloMito

Mitochondria phylogenomics analyses

- \$> Alignment of amino acids from each gene
- \$> Retrotanslation to nucleotides
- \$> Concatenation of gene sequences
- \$> Maximum likelihood analysis by supermatrix

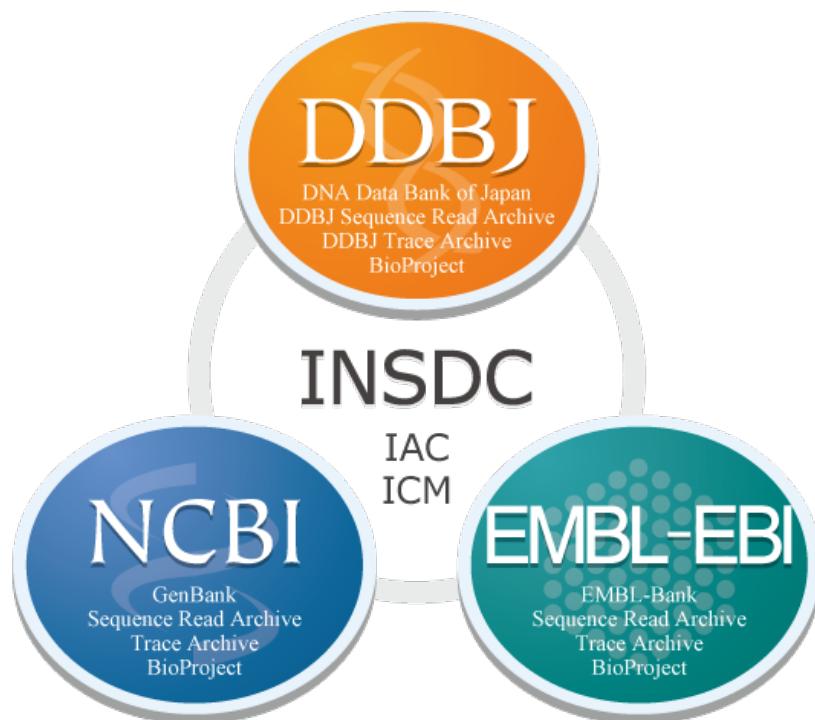
# Objetivos

- Montar, anotar e disponibilizar mitogenomas da família Formicidae
- Entender as relações filogenéticas do grupo por meio da análise de genes mitocondriais



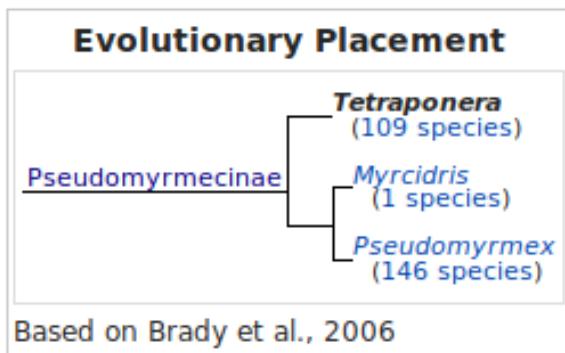
# Primeira etapa: Montagem

- Duas fontes de dados:
  - Bancos de dados públicos – Espécies com dados mas sem mitogenoma descrito
  - Nossas próprias amostras – Sequenciamento no INCA



# Bancos de dados públicos

- Grande quantidade de informação para o clado:
  - ≈ 3.91 Terabytes
- Todas as subfamílias
  - Exceto Martialinae
- Pseudomyrmecinae:
  - Subfamília pequena
  - Muitos dados



NCBI Entrez PubMed Nucleotide Protein Genome Taxonomy Browser

Search for Formicidae as complete name lock Go

Display 1 levels using filter: none

Nucleotide Nucleotide EST Nucleotide GSS Protein Structure Genome GEO Datasets UniGene PubMed Central Gene HomoloGene SRA Experiments GEO Profiles Protein Clusters Identical Protein Groups SPARCLE Bio Project Bio Sample Clone DB Genetic Testing Registry Host Viral Host Probe PubChem BioAssay

**Lineage (full):** cellular organisms; Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Bilateria; Protostomia; Hexapoda; Insecta; Dicondylia; Pterygota; Neoptera; Holometabola; Hymenoptera; Apocrita; Formicidae; Agroecomyrmecinae; Amblyoponinae; Aneuretinae; Apomyrminae; Dolichoderinae; Dorylinae; Ectatomminae; Formicinae; Heteroponerinae; Leptanillinae; Martialinae; Myrmeciinae; Myrmicinae; Nothomyrmeciinae; Paraponerinae; Ponerinae; Proceratiinae; Pseudomyrmecinae; unclassified Formicidae

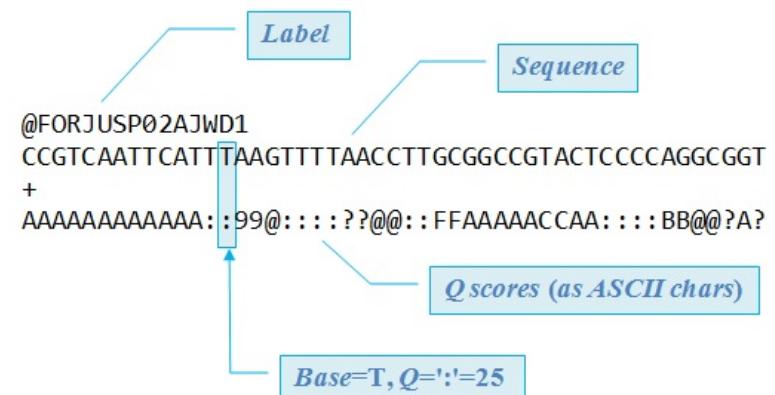
- [Formicidae](#) 3,368 Click on organism name to get more information.
  - [Agroecomyrmecinae](#) 1
  - [Amblyoponinae](#) 3
  - [Aneuretinae](#) 2
  - [Apomyrminae](#) 1
  - [Dolichoderinae](#) 201
  - [Dorylinae](#) 388
  - [Ectatomminae](#) 4
  - [Formicinae](#) 863
  - [Heteroponerinae](#) 2
  -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

# Obtenção dos dados

- Arquivos *.sra*
  - Utilizado pelos 3 bancos
  - Disponíveis para download direto
  - Precisa ser convertido (SRA Toolkit) para *.fastq*

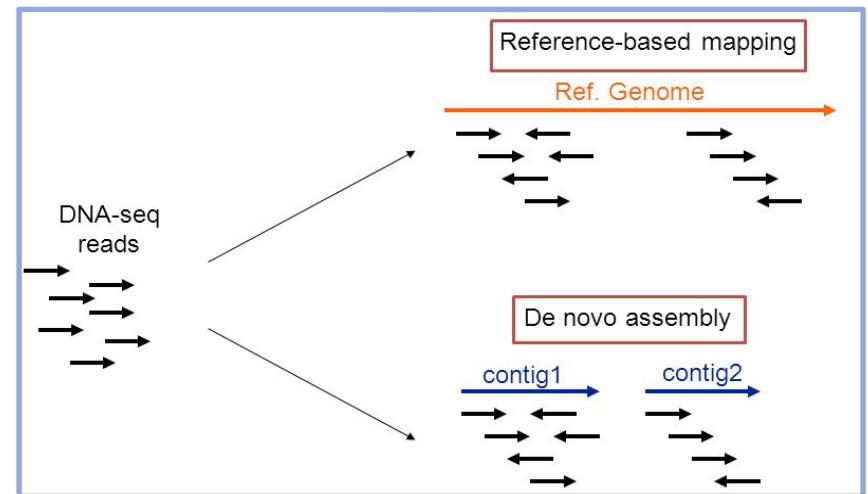
Study accession	Sample accession	Secondary sample accession	Experiment accession	Run accession	Tax ID	Scientific name	Instrument model	Library layout	FASTQ files (FTP)	FASTQ files (Galaxy)	Submitted files (FTP)	Submitted files (Galaxy)	NCBI SRA file (FTP)	NCBI SRA file (Galaxy)	CRAM Index files (FTP)	CRAM Index files (Galaxy)
PRJNA360290	SAMN06208930	SRS1901018	SRX2468701	SRR5150611	55425	Paraponera clavata	Illumina HiSeq 2500	PAIRED	File 1 File 2	File 1 File 2			File 1	File 1		

- Arquivos *.fastq*
  - Cabeçalho, sequência e qualidade
  - Aceitos pela maioria dos programas



# Montagem dos mitogenomas de Pseudomyrmecinae

- NOVOPlasty
  - Montagem de novo
  - Não gera alinhamento
  - Avisa quando circulariza
- Circularizou: montagem com dataset menor (MIRA/MITObim)
- Não circularizou: maior contig usado como referência no MITObim, com o dataset inteiro
- 14 mitogenomas completos
  - 12 *Pseudomyrmex* spp.
  - 2 *Tetraponera* spp.



# Montagem dos mitogenomas de Pseudomyrmecinae

- Circularizados por NOVOPlasty (12)
  - *Pseudomyrmex concolor*
  - *Pseudomyrmex dendroicus*
  - *Pseudomyrmex elongatus*
  - *Pseudomyrmex feralis*
  - *Pseudomyrmex ferrugineus*
  - *Pseudomyrmex flavidornis*
  - *Pseudomyrmex janzeni*
  - *Pseudomyrmex pallidus*
  - *Pseudomyrmex particeps*
  - *Pseudomyrmex peperi*
  - *Pseudomyrmex veneficus*
  - *Tetraponera rufonigra*
- Circularizados por MITObim (2)
  - *Pseudomyrmex gracilis*
  - *Tetraponera aethiops*

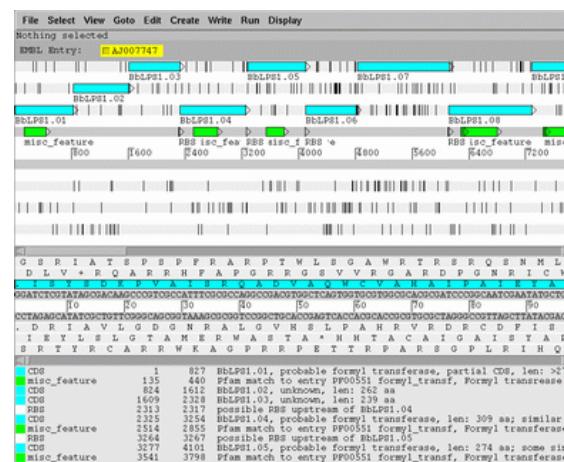
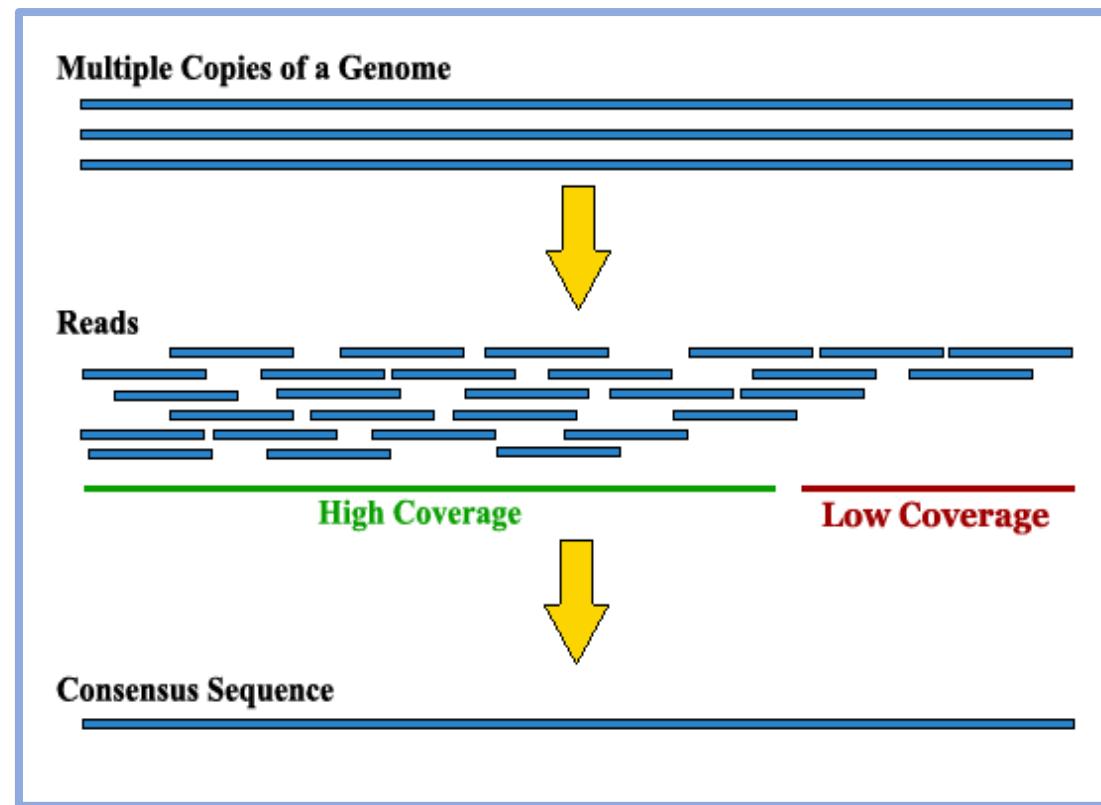
# MONTAGEM

- Checagem (TABLET)
    - Cobertura
      - Regiões de baixa cobertura anotadas
    - Circularização



- ANOTAÇÃO
    - Automática: MITOS Web Server
    - Manual: Artemis

■ tRNA gene ■ rRNA gene ■ protein coding gene



# Pseudomyrmecinae

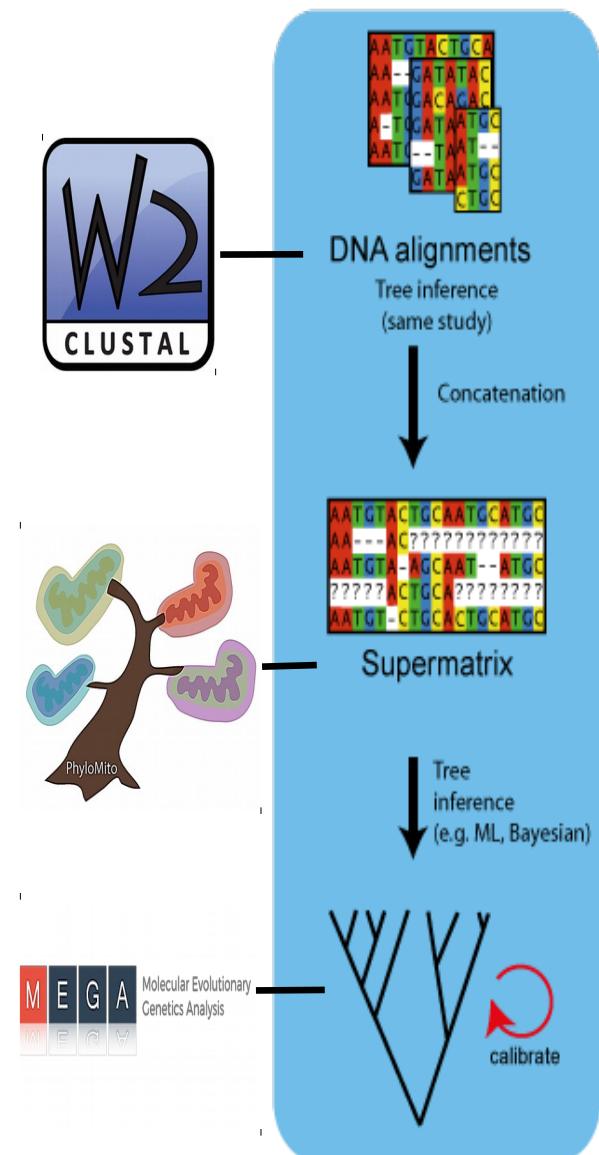
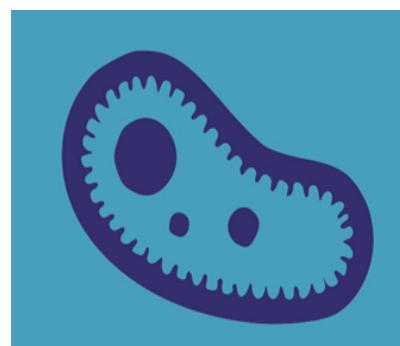
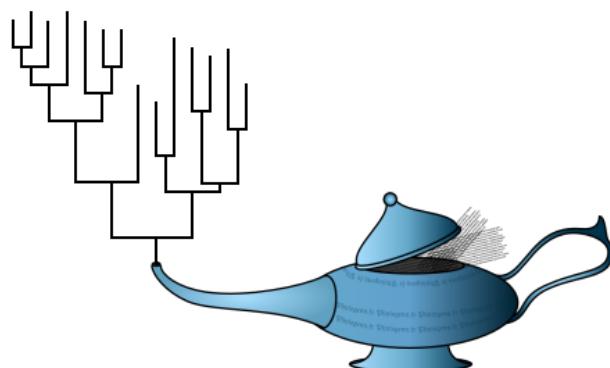


- Grande variação do tamanho da sequência:
  - 15704 (*Pseudomyrmex gracilis*) até 18835 (*Pseudomyrmex feralis*)
- Mesma sintenia da subfamília Dolichoderinae:
  - Clados evolutivamente próximos



# Perspectivas

- Submeter as sequências a um banco de dados público
- Realizar filogenômica e outras análises com mitogenomas da família Formicidae, incluindo os gerados nesse trabalho
- Escrever manuscrito relatando nossos resultados
- Realizar o mesmo processo para as outras subfamílias



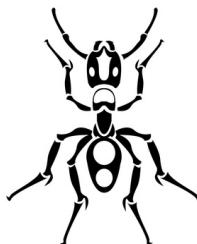
# PROJETO MITOGENOMAS

- Obter genomas mitocondriais usando dados públicos
- Descobrir a história evolutiva contada pela mitocôndria



# PROJETO MITOGENOMAS

- FORMIGAS



Stephanie

Juan

Vitor

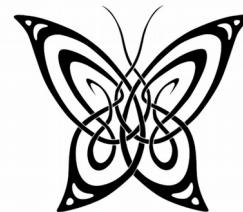
Pedro

- LAGARTOS



Gabriela

- BORBOLETAS



Vanessa

- TARTARUGAS



Ana

Deise

# OBRIGADO!

