

# Biodiversity and adaptations of CYP enzymes in the Amazon

## Loricariidae fishes

PI\_BR: Dr. Thiago E. Parente – FIOCRUZ, Rio de Janeiro, Brazil

PI\_US: Dr. Mark E. Hahn – WHOI, Woods Hole, MA, USA  
[parente@ensp.fiocruz.br](mailto:parente@ensp.fiocruz.br) & [mhahn@whoi.edu](mailto:mhahn@whoi.edu)

### Colaborators

Dr. John Stegeman – WHOI, USA  
 Dr. Francisco Paumgarten – FIOCRUZ, Brazil  
 Dr. Paulo Buckup – MN, UFRJ, Brazil  
 Dra. Carolina Furtado – INCA, Brazil  
 Dr. Adalberto Val – INPA, Brazil  
 Dr. Brian Haas – Broad Institute, USA

### STAFF

MSc. Daniel A. Moreira  
 Paula C.C. de Andrade (undergrad)  
 Maithê G.P. Magalhães (undergrad)



Hypancistrus zebra

## Abstract

Loricariidae is the most diverse family of the fish order Siluriformes<sup>1</sup>. Currently, there are >700 valid species and this number is still growing due to discovery of new species and the resolution of cryptic ones<sup>1,2,3</sup>. Loricariids inhabit a wide spectrum of fresh water environments, have diverse ecological habitats and show high endemism rate<sup>4</sup>. This astonishing taxonomic and ecological diversity must rely on genetic, molecular and biochemical adaptations. In fact, our group has reported peculiar adaptations in loricariids, which might increase their susceptibility to the toxic effects of pollutants<sup>5,6,7</sup>. Here, the molecular biodiversity of Loricariidae fish is explored using Next-Generation Sequencing (NGS) technologies. In total, 101 species has been collected; 34 of these will have their liver transcriptome sequenced. For one species, *Hypancistrus zebra*, the transcriptome of brain, heart, kidney, gill, intestine and gonad will also be sequenced. The transcriptome of three species has been sequenced; *Pterygoplichthys anisitsi*, *Ancistrus* spp. and *Corydoras nattereri*. The results for *P. anisitsi* are shown below. Briefly, 34204 genes with homologs in zebrafish were sequenced; 189 of these contained the complete coding sequence of proteins involved in mechanisms of defense against toxins. Great expansions were detected in subfamilies of Cytochromes P450 and Sulfotransferases.

References:  
 1. Covain & Fisch-Muller, 2007; Zootaxa 1462: 1-40  
 2. Ferraris, 2007; Zootaxa 1418: 1-628  
 3. Taphoni et al., 2013; Zootaxa 3641(4): 343-370  
 4. Covain & Fisch-Muller, 2007; Zootaxa 1462: 1-40  
 5. Parente et al., 2009; Comp. Biochem. Physiol. C 150: 252-260  
 6. Parente et al., 2011; Gene 489: 113-118  
 7. Parente et al., 2016; Aquatic Toxicology 154: 193-199

## Results

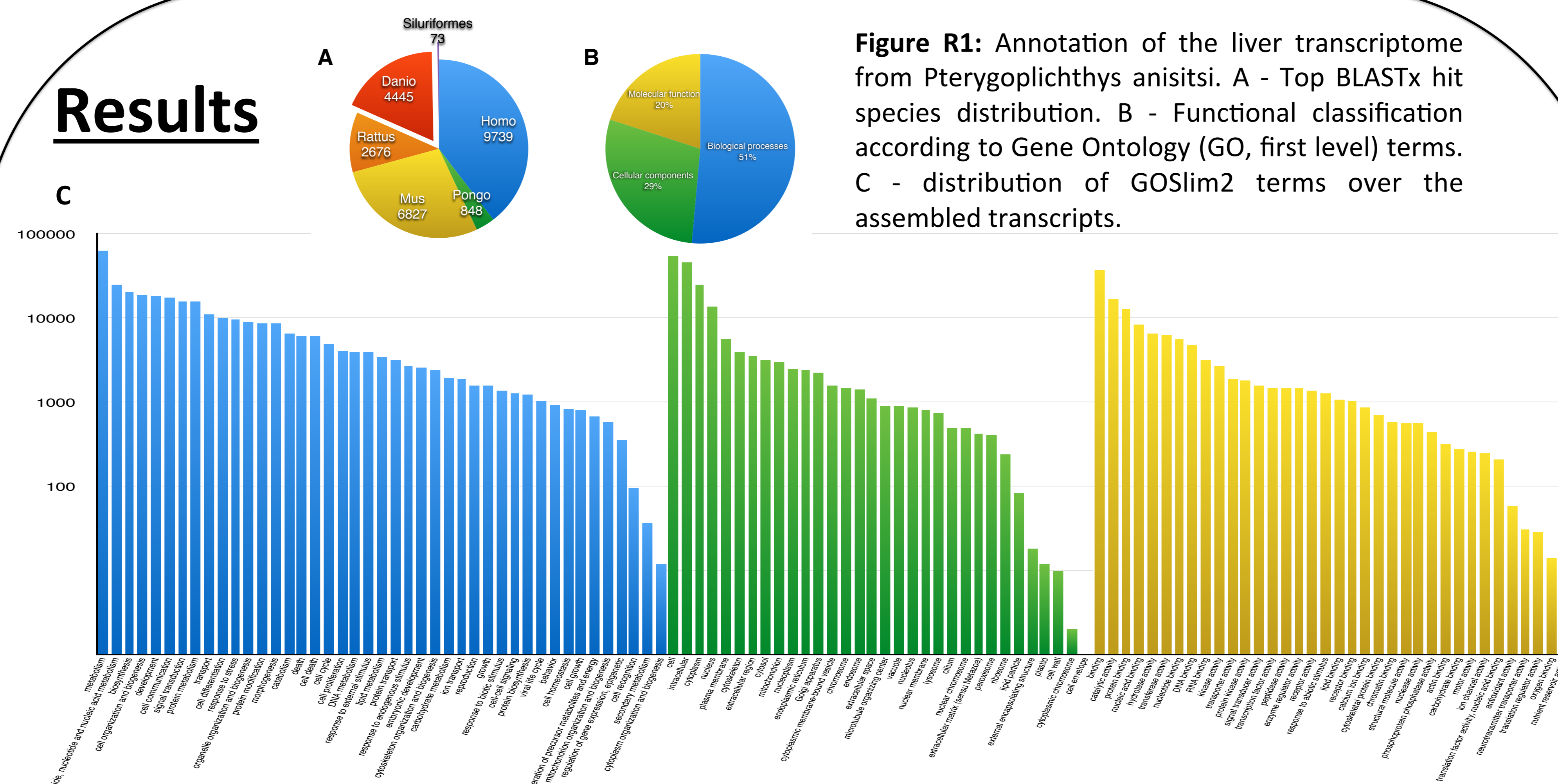
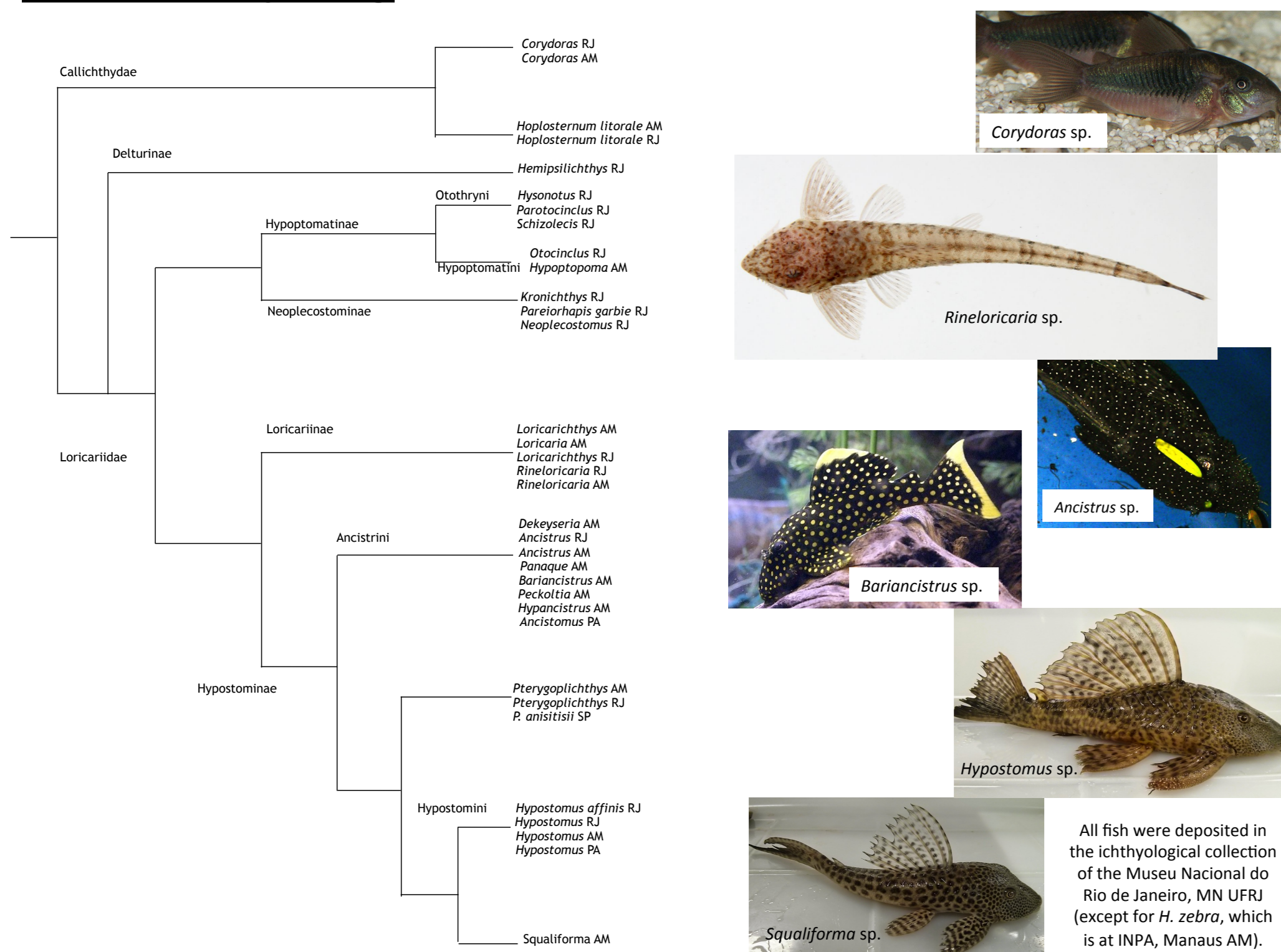


Figure R1: Annotation of the liver transcriptome from *Pterygoplichthys anisitsi*. A - Top BLASTx hit species distribution. B - Functional classification according to Gene Ontology (GO, first level) terms. C - distribution of GOslim2 terms over the assembled transcripts.

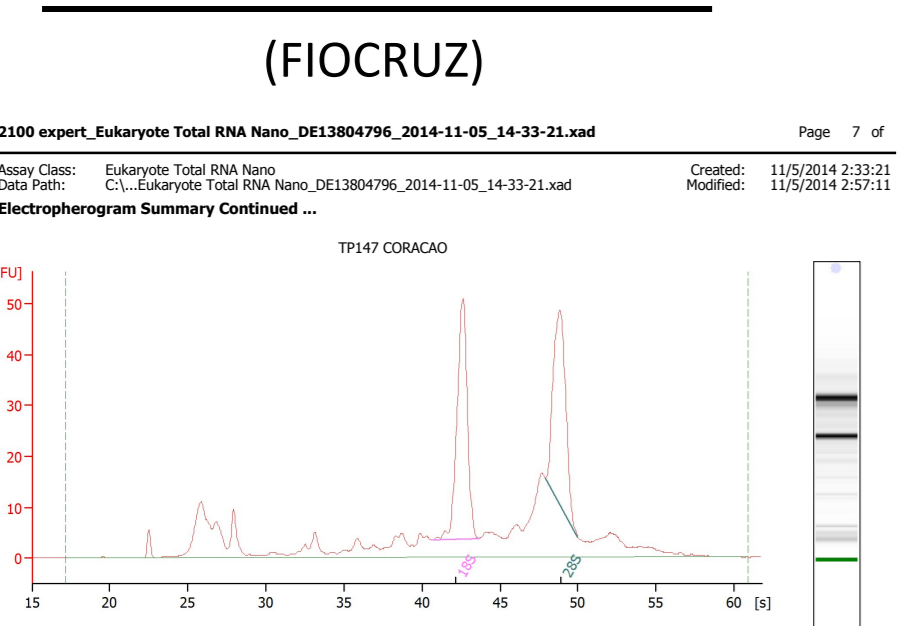
## Material & Methods

### Fish sampling (FIOCRUZ, INPA and MN UFRJ)

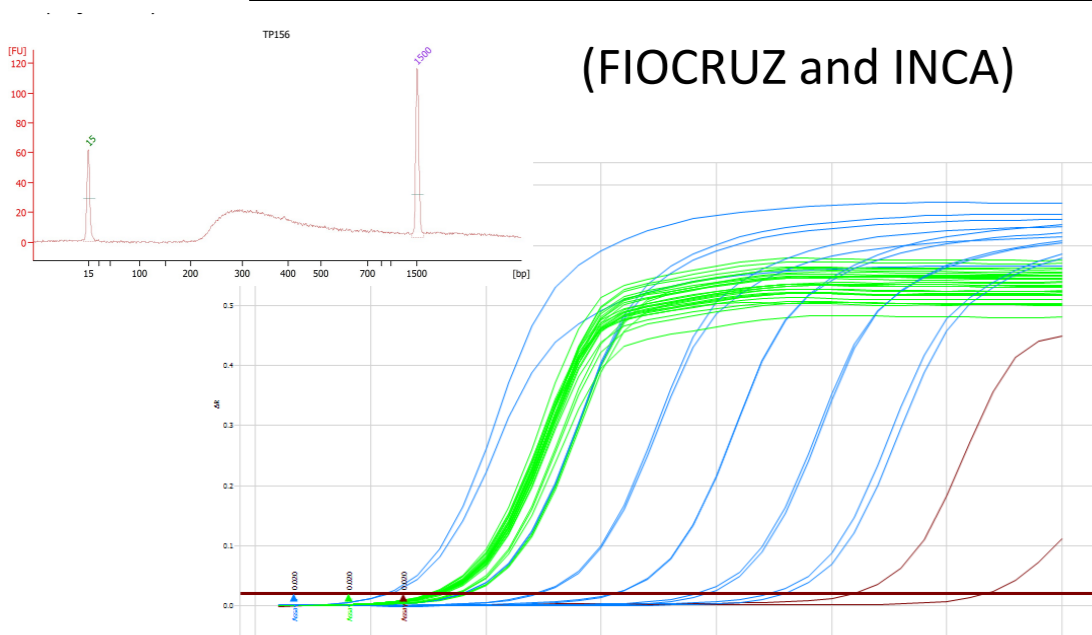


All fish were deposited in the ichthyological collection of the Museu Nacional do Rio de Janeiro, MN UFRJ (except for *H. zebra*, which is at INPA, Manaus AM).

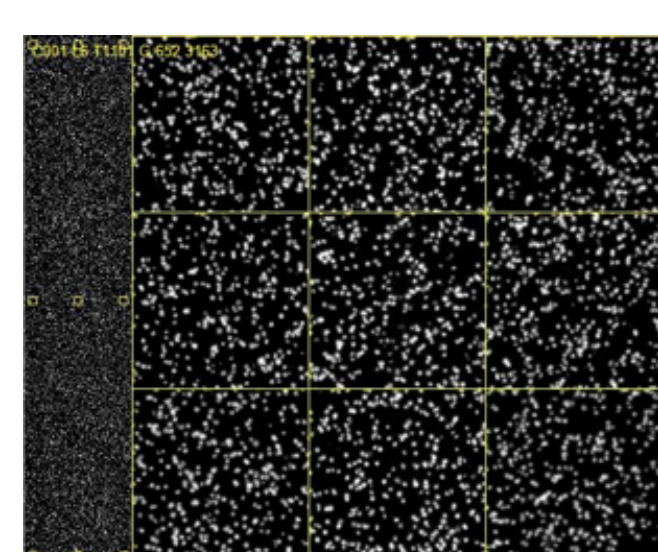
### RNA extraction (FIOCRUZ)



### Library preparation (FIOCRUZ and INCA)



### Illumina sequencing (INCA)



### Bioinformatic

Table M1: Summary of the three transcriptomes sequenced so far.

| Species                          | Reads (paired > 30) | Transcripts | BLASTX hits against UNIPROT | Human | Zebrafish |
|----------------------------------|---------------------|-------------|-----------------------------|-------|-----------|
| <i>Pterygoplichthys anisitsi</i> | 59118142            | 66670       | 30354                       | 34204 |           |
| <i>Ancistrus</i> sp.             | 60170745            | 67883       | 29891                       | 33953 |           |
| <i>Corydoras nattereri</i>       | 55162726            | 52565       | 22827                       | 28510 |           |

Tools used: Trinity, BLAST, UniProt, eggNOG, iGV, and others.

Figure R2: Mitochondrial genome of *Pterygoplichthys* spp. A - Circular representation of *P. disjunctivus* mtDNA (GI: 339506171). Red circles indicate the tRNA not sequenced in *P. anisitsi* mtDNA and red arrows the approximate region of 10, 24 and 31 nucleotides. B - Linear representation of the annotated mitochondrial genome of *P. anisitsi*.

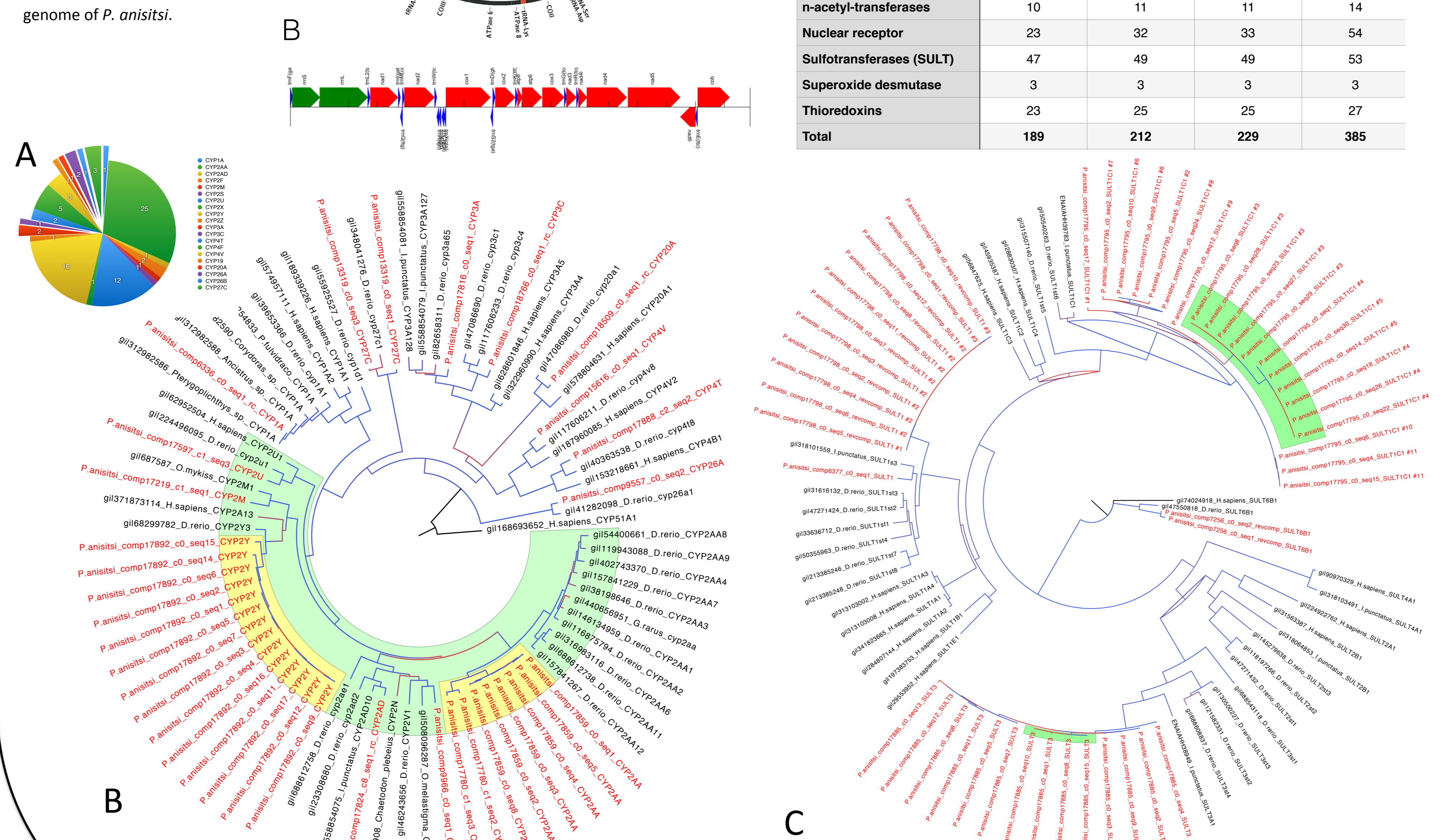


Figure R3: Distribution of cytochromes P450 (CYP) into subfamilies (A); and the maximum likelihood phylogenetic tree of the sequenced CYP (B) and Sulfotransferases (SULT) genes (C).

Table R1: Number of sequenced components in the hepatic transcriptome of *P. anisitsi* with complete coding sequence (CDS), >75% of the CDS, >50% of the CDS and the total number of contigs for each defense gene family.

| Gene Family                | Full length CDS | >75% CDS coverage | >50% CDS coverage | Contigs    |
|----------------------------|-----------------|-------------------|-------------------|------------|
| AHR & ARNT                 | 3               | 3                 | 3                 | 9          |
| Aldo Keto Reductase        | 5               | 5                 | 5                 | 5          |
| ATP Binding Cassette (ABC) | 13              | 19                | 25                | 91         |
| Basic leucine zipper       | 3               | 3                 | 5                 | 8          |
| Catalase                   | 1               | 1                 | 1                 | 1          |
| Cytochrome P450            | 34              | 35                | 39                | 82         |
| Epoxide hidroxilase        | 2               | 2                 | 2                 | 2          |
| Glucuronosyltransferase    | 8               | 8                 | 10                | 15         |
| Glutathione Peroxidase     | 6               | 6                 | 8                 | 11         |
| Glutathione-S-transferase  | 8               | 10                | 10                | 10         |
| n-acetyl-transferases      | 10              | 11                | 11                | 14         |
| Nuclear receptor           | 23              | 32                | 33                | 54         |
| Sulfotransferases (SULT)   | 47              | 49                | 49                | 53         |
| Superoxide desmutase       | 3               | 3                 | 3                 | 3          |
| Thioredoxins               | 23              | 25                | 25                | 27         |
| <b>Total</b>               | <b>189</b>      | <b>212</b>        | <b>229</b>        | <b>385</b> |

## Tangential project #1

### Mitogenomes assembled from transcriptome

We sequenced the transcriptome of three fish from the genus *Ancistrus* (Loricariidae, Siluriformes) using as start material total RNA isolated from the liver. The transcriptome data were used to assemble the mitogenome of each fish with 92%, 95% and 99% of the full length of their closest related species with a sequenced mitogenome. Taken the sequences of the three fish together, all the 13 protein-coding genes, two ribosomal RNAs, 22 tRNAs and the D-loop known in the mitogenomes of vertebrates were sequenced. The use of transcriptomic data also allowed the clear observation of the punctuation pattern of mtRNA editing, to analyze the transcriptional profile of mtRNA, and to detect heteroplasmic sites.

Table TP1: Summary of the transcriptome and mitogenome data for the three fish (*Ancistrus* spp.).

|                            | <i>Ancistrus</i> sp. #1 | <i>Ancistrus</i> sp. #2a | <i>Ancistrus</i> sp. #2b |
|----------------------------|-------------------------|--------------------------|--------------------------|
| RNA Integrity Number - RIN | 8.2                     | 7.4                      | >7.00                    |
| Library insert size (bp)   | 230-800                 | 268-792                  | 285-370                  |
| Reads after QC             | 43502597                | 53961751                 | 60170745                 |
| Transcripts                |                         |                          |                          |
| total                      | 67098                   | 63847                    | 67883                    |
| with BLASTx hit            | 57110                   | 51886                    | 53953                    |
| for mitogenome             | 13                      | 12                       | 7                        |
| mtRNA reads (%)            | 2.6                     | 1.8                      | 0.8                      |
| Mitogenome coverage (%)    | 99.2                    | 92.5                     | 94.7                     |
| Heteroplasmic sites        | 44                      | 46                       | 41                       |

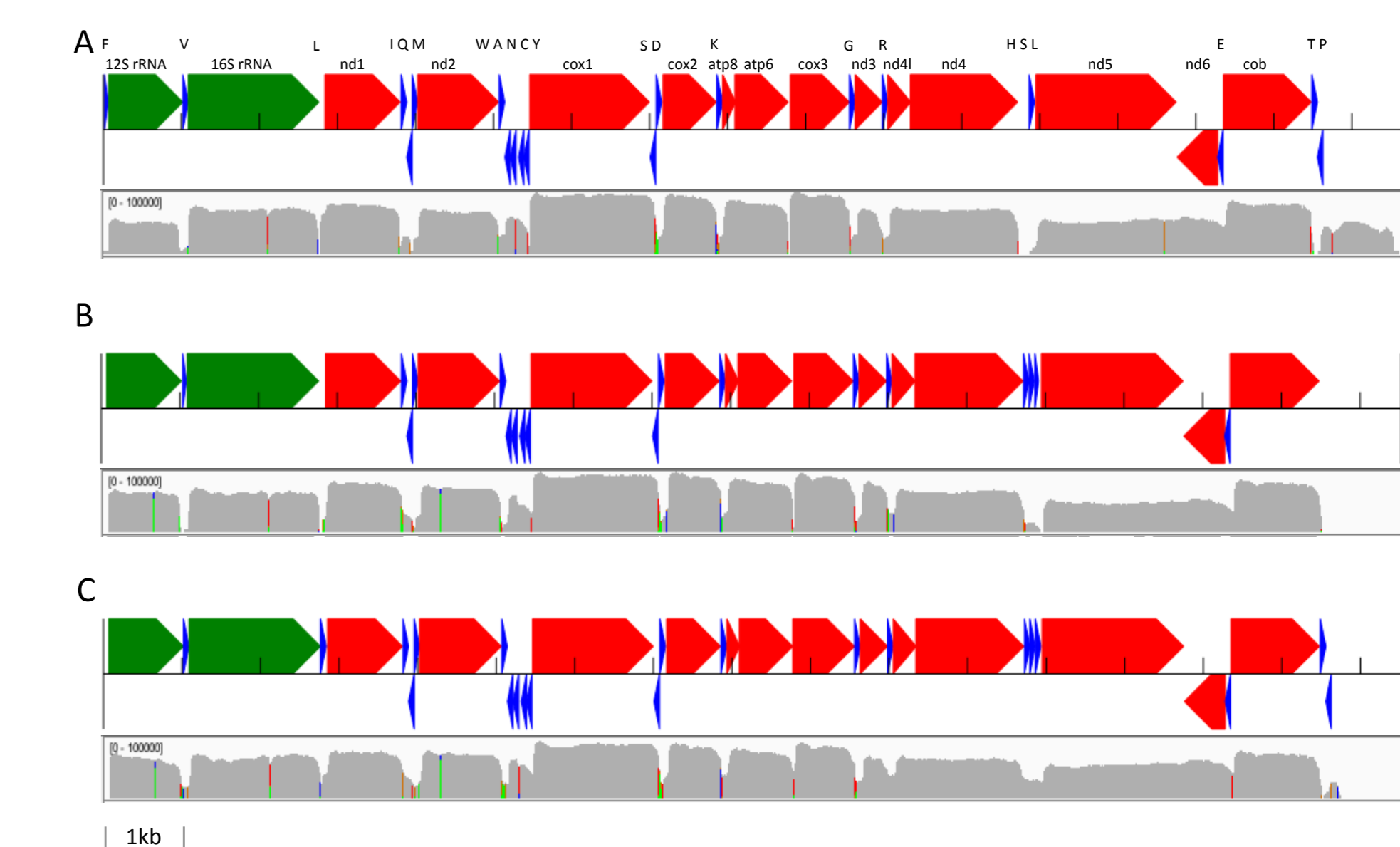
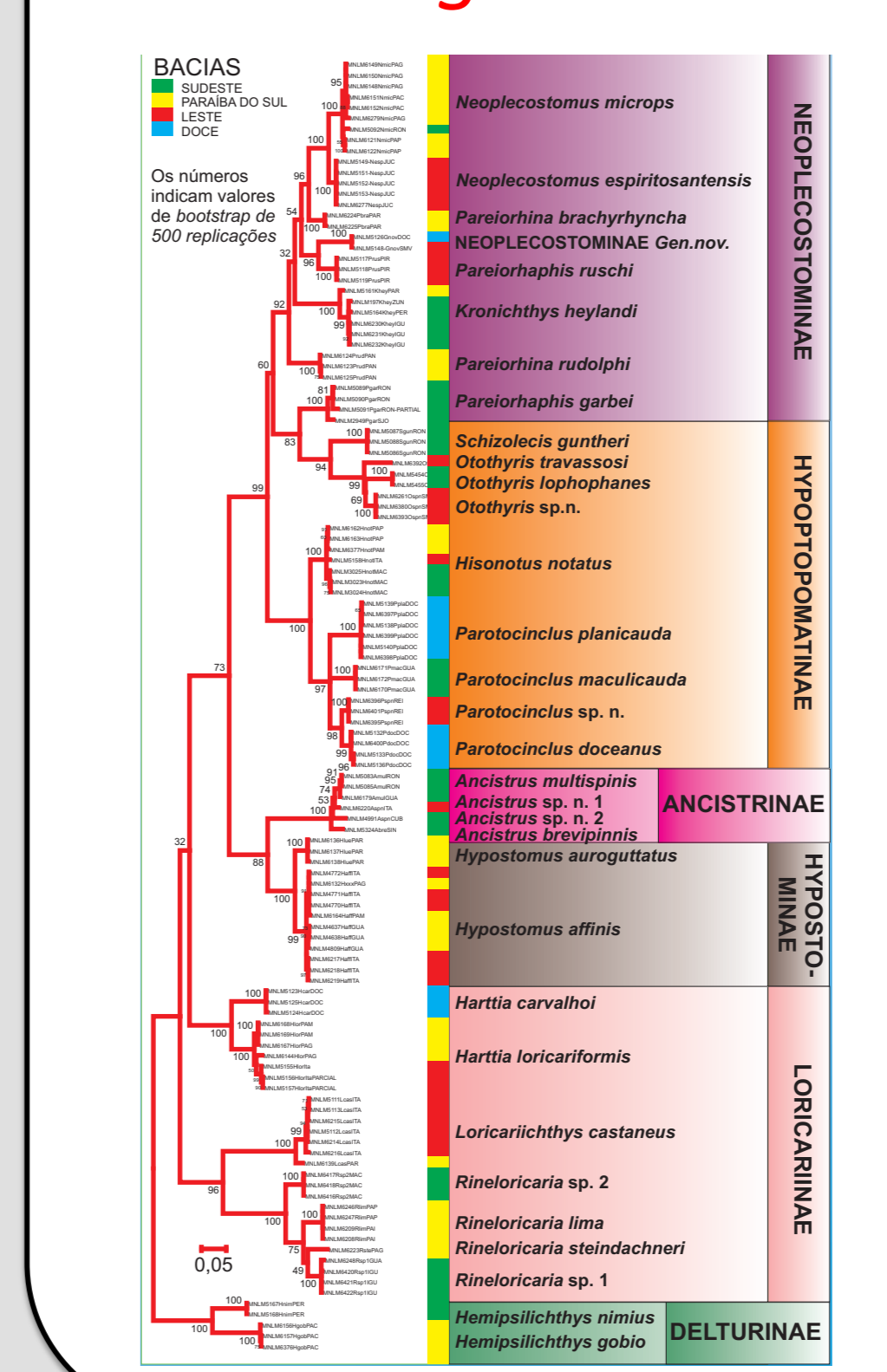


Figure TP1: The assembled mitogenomes of *Ancistrus* sp. #1 (A), *Ancistrus* sp. #2a (B) and *Ancistrus* sp. #2b (C). The number of supporting reads along the sequence is shown in a logarithmic scale and below the schematic view of each mitogenome. Heteroplasmic sites are highlighted with a different color on the graphic of supporting reads. Each feature of the mitogenomes is named on the top of the figure. The tRNAs are named using the one-letter code of the amino acid they transport.

## Tangential project #2

### Barcoding loricariids



## Yet to be done

- Sequence the transcriptome of 31 species
- Assemble the 31 transcriptomes
- Analyze CYPs and AHR genes in the 31 species
- Amplify and sequence CYP1A and AHR in the others sampled species
- Exposure of selected species to chemicals
- Analyze alterations in gene transcription

## Developmental impact

- Current:**
- Database for molecular ID of Loricariidae
  - ID of new and cryptic species and genera
  - ID of new genes
  - Training human resources
- Future:**
- Enlarge DB for molecular ID of Loricariidae
  - Evaluate loricariids responses to pollutants
  - Knowledge to conservation
  - Support BR polices to preserve biodiversity

## Publications

- Paper published:**
- Parente et al., 2015; Vigilância Sanitária em Debate 3(11): 88-93
  - Felício; Parente et al., 2015; Ecotoxicol. Environ. Safety 115: 26-32
  - Parente et al., 2014; Encontro Brasileiro de Genética
  - Buckup et al., 2015; Encontro Brasileiro de Ictiologia
- Chapter published:**
- Parente & Hauser-Davis, 2013; In: Pollution and fish health in tropical ecosystems; CRC Press
- Abstract and posters in conferences:**
- Parente et al., 2014; P450 Biodiversity and Biotechnology, Japan
  - Magalhães de Andrade et al., 2014; Congresso Brasileiro de Genética
  - Buckup et al., 2015; Encontro Brasileiro de Ictiologia
- Manuscripts in preparation:**
- Parente et al.; The transcriptome and diversity of defense genes in *P. anisitsi*.
  - Moreira, Furtado & Parente; Constructing mitogenomes from transcriptome
  - Moreira et al.; The mitogenome of *Corydoras nattereri*.
  - Buckup et al.; Biogeography of Loricariidae from Rio de Janeiro state, Brazil

## Acknowledgements

