



## Aquaculture 2019 - Meeting Abstract

### GENOME SEQUENCING AND DE NOVO ASSEMBLY OF TAMBAQUI *Colossoma macropomum*

Alexandre Rodrigues Caetano\*, Patrícia Ianella, Eduardo Sousa Varela, Samuel R. Paiva, Francisco Pereira Lobo, Leandro Carrijo Cintra, Adhemar Zerlotini Neto, Michel E.B. Yamagishi.

Embrapa Recursos Genéticos e Biotecnologia  
Brasília-DF, 71540135  
Brazil  
alexandre.caetano@embrapa.br

The South American freshwater fish Tambaqui (*Colossoma macropomum*) is the most important native aquaculture species in Brazil. Current production levels are above 150,000 mt per year. Recent efforts to domesticate, breed and raise the species in aquaculture systems has led to significant increases in production (>10-fold) over the last decade, and development of new technologies based on genomic tools and information could help further increase productivity and production growth rates. A Tambaqui genome assembly was generated with data produced from shotgun libraries from two different insert sizes and mate-pair libraries with four different sizes sequenced (2x150bps) with a HiSeq2000 platform. A total of 124.8Gbp quality-filtered nucleotides were sequenced which amount to 85x mean genome coverage, considering previously published information (C-value = 1.5pg = 1.467Gbp). Initial sequence assembly was performed with SOAPdenovo and generated 8.920 scaffolds spanning 1.54 Gbps (N50 scaffold length: 2,041kb, L50 scaffold count: 162). Gene model predictions with MAKER2 using as extrinsic evidence protein and EST data from phylogenetically related taxa found 23,632 genes, ~20,000 shared with *A. mexicanus*. CEGMA analysis (prior to gap filling) detected 85% core genes completely assembled and 9% partially assembled. Tambaqui assembly scaffolds ordering and orientation was carried out using ALLMAPS software based on the available linkage map (LM) (Nunes et al. 2017). A total of 1440 scaffolds were placed in the map with 27 linkage groups (n=x=27), totaling 1,508,696,562bp and 94% (6,788) of all mapped markers, with an average of 4.7 markers per megabase. Of the 6,788 mapped markers, ~93% were anchored, 84.9% in concordance with marker orientation, and 7.1% markers were unplaced. Obtained results reveal a high-quality NGS *de novo* Tambaqui genome assemble, providing a valuable tool for future genomic studies and the use of genomic tools for broodstock management and assisted genetic evaluations and breeding.