



## **Elaeis oleifera genome draft – genomics of american oil palm**

*Richardson Silva Lima, Francisco Pereira Lobo, Marcelo Falsarella Carazzolle, Gustavo Gilson Lacerda Costa, Emilio Luiz de Camargo Rodrigues, Alexandre Alonso Alves, Michel Eduardo Beleza Yamagishi, Manoel Teixeira Souza Junior, Eduardo Fernandes Formighieri*

Embrapa Agroenergy, Embrapa Agricultural Informatics, Laboratory of Genomics and Expression, Laboratory of Genomics and Expression, Embrapa Agroenergy, Embrapa Agroenergy, Embrapa Agricultural Informatics, Embrapa Agroenergy, Embrapa Agroenergy

---

The palm oil and palm kernel, extracted from Oil palm (*Elaeis guineensis*), represented around 36.7% of world vegetable oil production in 2009. The Oil palm productivity is around 10 times higher than Soybean's. The main producers and exporters are Indonesia and Malaysia (87% in 2009) and the main consumers are China and India. Over the past 20 years, the share of palm oil market has grown two-folds, with a strong increase in global palm oil trade. Palm oil can be used in food industry, to produce biodiesel, paints, antioxidants, cosmetics, solvents etc. Here in Brazil it is an important regional food, the well-known "óleo de dendê", and also presents high potential as renewable source of biodiesel. A condition called bud rot is causing a severe decrease in oil palm production worldwide, but an American species (*Elaeis oleifera*, or "caiaué") from the same genus is tolerant to this abnormality. The hybrid originated from the intercrossing of both species is viable, presents good productivity, better oil quality and is tolerant to bud rot. In the Northern region of Brazil, which concentrates most of the oil palm production, are located two Embrapa Research Centers which are involved in *Elaeis* research. One of these centers holds the Active Germplasm Collection of the species. Recently, Embrapa released an interspecific hybrid called "Manicoré". In that scenario, Embrapa Agroenergy is making efforts to improve bioenergy generation and accelerate selection of superior genotypes, through NGS and large-scale genotyping technologies. Some colleagues have already estimated of *Elaeis* species genome size using flow cytometry, and in parallel, we sequenced one lane of Illumina HiSeq 2000 for three *E. oleifera* and one *E. guineensis* accessions, for initial comparison. The technology and inserts size were chosen considering the AllPaths-LG approach. Next step included more sequencing for chosen accession (Manicoré) using both Paired-Ends and Mate-Pairs libraries from different insert sizes (180 and 800 bp for PEs, and 3, 5 and 10K for MPs). In this step, we tested correction with SOAPdenovo correction tool, quality control with fastx\_toolkit and some different assemblers and parameters (such as AllPaths-LG and SOAPdenovo). Recently we discussed alternatives for step three, and are requesting quotation for the final decision (at this time, involving different sequencing technology, with long inserts and/or long reads), and by 2013 we intend to present the genome draft of *Elaeis oleifera*. As future work, we will perform a semiautomatic annotation of structures (genes, tRNAs, etc.), setup the information in a genome browser and submit a manuscript for publication. Supported by: FINEP, CNPq and Embrapa.

**Keywords:** American Oil Palm, *Elaeis oleifera* genome draft, NGS, AllPaths-LG approach

**Concentration area:** Genomics Evolution

---

