

# Reconstructing the Whole Mitochondrial DNA (mtDNA) from Nuclear Genome

Giovanni Marques de Castro, Adhemar Zerlotini Neto, Michel Eduardo Beleza Yamagishi

*Embrapa Informática Agropecuária*

## Abstract

In several eukaryotic organisms, the nuclear genome has several partial copies of the mitochondrial DNA (mtDNA). These copies are called NUMTs (NUclear MiTochondrial DNA) and they have been known since 1967 when the first evidence of them were reported in the mouse nuclear genome. Despite almost fifty years have passed, the reason of their very existence remains controversial. However, their presence has been confirmed in an increasing number of genomes. The NUMTs could be only another DNA idiosyncrasy, but they actually represent a serious issue for important application such as genome bar coding. There are many open questions about them. A practical one could be stated as: is it possible to reconstruct the whole mtDNA from the NUMTs, and how similar this hypothetical sequence would be to the actual mtDNA. In order to address this question, we have taken advantage of PacBio reads derived from a *Bos indicus* mtDNA-free sperm sample, and after a mapping procedure that sorted out nuclear from NUMT-like reads, the Mira assembler was applied to get a sequence that has 97% of identity to the actual *Bos indicus* mtDNA. Although our bovine experiment can not be straightforward extrapolated to other organisms without further investigations, it has helped us to answer the aforementioned question, and might suggest that DNA samples, where the mtDNA is lacking or so damaged to the point of preventing its assembly, can, nevertheless, deliver enough information to reconstruct a sequence that resembles the actual mtDNA. For instance, ancient DNA derived from fossils may be an example of interest.