Reverse transcriptases drive the transition from the RNA world to the DNA world

ADRIEN BONIFACE[1]*, CATHERINE LAROSE[1], TIMOTHY M. VOGEL[1]

[1]Environmental Microbial Genomics, Laboratoire Ampère, Ecole Centrale de Lyon, Université de Lyon, 6910 Ecully, France (*correspondence : adrien.boniface@univ-lyon1.fr)

The first life forms were probably different from what we now consider as the basic unit of life, the cell. The complexity of the current cellular system is such that it is unlikely to have emerged spontaneously from abiotic chemical reactions, therefore the first life forms were probably considerably simpler. One hypothesis is that life originated in an "RNA World" with RNA molecules (ribozymes) carrying out the conservation and replication of the genetic information, and the catalysis of metabolic reactions. Life then evolved into a modern "DNA World" where the functions are divided between the different classes of biological molecules. At the transition between the RNA world and the DNA world, a process to convert the RNA-encoded genomes into DNA-encoded genomes, called reverse transcription, must have occurred [1]. This reaction is catalyzed by proteins called reverse transcriptases (RTs), which were first discovered in retroviruses. Eukaryotic genomes are rich in integrated retroviral sequences, making them rich in RT genes, but archaeal and bacterial genomes have been under explored for RT genes. If modern RTs are inherited from proteins that were essential for the RNA-to-DNA-genome transition, then homologous genes (or sequences) should be found in all modern genomes as a vestige of this fundamental transformation of life, even if their function might have changed over time. Therefore, this study characterized the diversity of modern RTs by finding genes homologous to known RTs in organisms as diverse as possible, based on the available fully sequenced genomes and metagenomic datasets. The first step was to compare known RTs to identify conserved regions of the genes susceptible to retain phylogenetic information across a wide diversity of organisms and to calibrate the stringency of the analysis. A second step will consist of a phylogenetic analysis in order to determine to what extant the diversity of RT-homologous genes results from vertical inheritance or horizontal gene transfer. These results will help understand what must have been a fundamental transformation in the history of life.

[1] Forterre (2005) *Biochimie* **87,** 793–803.