Xenopolymerases and nucleic acid evolution. Andrew D. Ellington,[1] Jimmy Gollihar[1], and Jared Ellefson[1]
[1] Center for Systems and Synthetic Biology, University of Texas at Austin, Austin, TX, 2500 Speedway, 78712, USA

The evolution of nucleic acids is widely thought to be a defining moment in the origin and early evolution of life; however the accumulation of significant functional information in a nucleic acid-based organisms would almost certainly required polymerase functionality, whether RNA or protein. Given the RNA world hypothesis, it is therefore odd that there are no deep lineages for reverse transcriptases (RTs); rather, most reverse transcriptases currently known are for the most part offshoots of a more ancient RNA polymerase lineage and are found in viruses or selfish genetic elements. We were curious whether this observation was due to history or function; was it difficult or impossible to first invent RT activity, or were ancient RT lineages lost during early evolution?

Starting from a thermostable archael DNA polymerase (KOD), we were able to use directed evolution to readily acquire reverse transcriptase activity. The resultant enzyme, RTX, is the first known reverse transcriptase associated with DNA polymerase phylogeny, and suggests that it would have been readily possible to traverse between RNA, DNA, and reverse transcriptase lineages. Moreover, RTX is an error-correcting reverse transcriptase, capable of amplifying both RNA and DNA. This makes it inordinately useful for biotechnology applications, but also raises the interesting question of why there are no long RNA genomes, as it should be possible to replicate them via DNA with suitable error correction. The answer may lie, of course, in the inherent instability of RNA itself relative to DNA, but it is interesting to note that such a choice may have been set in living systems even prior to the evolution of robust polymerase functionality.

Further challenging paradigms as to what nucleic acids may have come first, and why, RTX has been elaborated to fully incorporate 2'-OMe nucleotides into nucleic acids. Such a modification would likely have been prebiotically accessible, perhaps more readily than a hydrogen for hydroxyl substitution at the 2' position. Thus, it is again an open question as to why life did not proceed from 2'-OH to 2'-OMe, or if it did, why it continued along the path to 2'-H?