Enzyme engineering through evolution: Thermostable recombinant group II intron reverse transcriptases provide new tools for RNA research and biotechnology

KATHLEEN COLLINS[1,3] and TIMOTHY W. NILSEN[2,3]

[1]Department of Molecular and Cell Biology, University of California at Berkeley, Berkeley, California 94720, USA
[2]Center for RNA Molecular Biology, Case Western Reserve University, Cleveland, Ohio 44106, USA

ABSTRACT

Current investigation of RNA transcriptomes relies heavily on the use of retroviral reverse transcriptases. It is well known that these enzymes have many limitations because of their intrinsic properties. This commentary highlights the recent biochemical characterization of a new family of reverse transcriptases, those encoded by group II intron retrohoming elements. The novel properties of these enzymes endow them with the potential to revolutionize how we approach RNA analyses.

Keywords: RNA-seq; noncoding RNA; reverse transcriptase; template switching

In biology, even hallowed rules like the central dogma have exceptions. Parallel to evolutionary diversification of the DNA-templated polymerases necessary for genome replication and expression, RNA-templated DNA and RNA polymerases are also evolutionarily diverse and widespread (Ng et al. 2008; Finnegan 2012). Among the polymerases that counter the central dogma flow, the best known are the reverse transcriptases (RTs). The initial discoveries of retrovirally encoded RT activity and function were recognized with the 1975 Nobel Prize in Physiology or Medicine. Recombinant retroviral RTs and their laboratory derivatives have been crucial research tools for molecular biologists for decades.

RTs other than retroviral family members are less well understood in structure or enzymology and have not yet been harnessed for commercial or medical diagnostic applications. One example is telomerase RT (Blackburn and Collins 2011), the biological significance of which was also recognized by the Nobel Prize in Physiology or Medicine in 2009. In addition to the retroviral and cellular RTs that maintain their respective genomes, RTs are also exploited by selfish DNA elements for their perpetuation. Genome-embedded retrotransposons encode RTs required for element mobility. In humans, non-LTR retrotransposon RTs copy polyadenosine-tailed templates to insert complementary DNA (cDNA) at nonspecific target sites, often with 5′ truncation (Finnegan 2012). In contrast, mobile group II introns encode RTs that copy the intron RNA to insert a precise full-length cDNA into an intronless allele of the host gene (Lambowitz and Zimmerly 2004). This process, termed retrohoming, must maintain the host gene exon sequences as precisely as the intron sequence to allow functional transcript production by intron splicing. Genome projects have unearthed hundreds of group II introns, mainly in prokaryotes and the organellar genomes of fungi and plants (Lambowitz and Zimmerly 2011).

Group II intron RTs synthesize long cDNAs with high fidelity, but they have remained untapped as a source of RTs for biotechnology applications. Group II intron RTs associate tightly with their coevolved intron RNA templates, like non-LTR retroelement RTs and telomerase. This stymies the potential commercial applications of these RTs in copying heterologous RNA templates. Furthermore, due to a physiological dependence on protein and RNA cofolding, RTs other than retroviral family members have not been amenable to recombinant protein expression. After many years of groundwork, the Lambowitz group reports in *RNA* a landmark study attaining robust high-level recombinant expression of group II intron RTs, focusing on thermostable enzymes from the bacterial thermophiles *Thermosynechococcus elongatus* and *Geobacillus stearothermophilus* (Mohr et al. 2013). The methods used to produce these thermostable group II intron RTs are generalizable to the group II intron RTs from other evolutionary branches of the tree of life.

As would be predicted from their biological function, group II intron RTs have higher processivity, fidelity, and thermostability than retroviral RTs (Mohr et al. 2013). Retroviral RTs copy their own genomes, or heterologous templates, with high error rates relative to cellular DNA polymerases as an

[3]**Corresponding authors**
E-mail twn@case.edu
E-mail kcollins@berkeley.edu
Article published online ahead of print. Article and publication date are at http://www.rnajournal.org/cgi/doi/10.1261/rna.040451.113.

evolutionarily optimal mechanism for virus evolution (Duffy et al. 2008). In contrast, retrohoming group II intron RTs must synthesize an accurate cDNA copy of the intron RNA, which can be quite long and also must be folded into the stable secondary and tertiary structures that support self-splicing. Thus, by evolutionary selection, group II intron RTs must have high processivity of cDNA synthesis on challenging template RNAs and also have high cDNA synthesis fidelity.

Even without optimization with rounds of structure-guided design and genetic selection, group II intron RTs have major advantages for RT applications, including quantitative RT-PCR and next-generation RNA sequencing. For example, thermostability of the group II intron RTs allows cDNA synthesis at temperatures that denature template RNA secondary structure, and their high processivity enables the synthesis of long cDNAs with minimal background from premature termination (Mohr et al. 2013). Perhaps the most striking and unanticipated property of the group II intron RTs is their ability to perform end-to-end template switching. Mohr and colleagues (Mohr et al. 2013) demonstrate that a DNA/RNA hybrid "primer" with a single 3′ nucleotide DNA overhang can, with high efficiency, form a base pair with the 3′ terminus of any target RNA. The DNA component of the "primer" is then extended by group II intron RT synthesis across the entire target RNA template. This method has many new applications and advantages over retroviral RTs for cDNA synthesis and library construction. For example, this approach eliminates the cumbersome, often inefficient, highly sequence-biased RNA ligation step common to many library construction kits. Moreover, this approach can be used to obtain cDNAs for otherwise "unclonable" RNAs with a modified or structured 3′ end. It will be of considerable interest to investigate what types of RNAs have remained hidden from identification simply because they are not captured with current cloning methodologies.

Remarkable new feats of cDNA synthesis are described in this initial *RNA* article, but, of course, myriad others are possible to envision. Based on our experiences chairing recent National Institutes of Health Common Fund study sections, the surging interest in extracellular RNA profiling as a diagnostic would benefit from less-biased cloning of structured RNAs as well as processed short microRNAs in serum samples. Also, unlike the retroviral RTs, group II intron RTs are selective for RNA versus DNA templates, a specificity that will greatly reduce complications arising from DNA contamination of RNA preparations for cDNA library construction. Starting from cross-linked chromatin fragments, group II intron RT cDNAs could be joined to nearby genomic DNA to obtain an unbiased genome-wide map of which noncoding RNAs are associated with which DNA loci.

Clearly, applications using recombinant group II intron RTs have enormous promise for the improvement of human health. It is notable that, against the wisdom of current funding trends, the National Institutes of Health made possible these applications by funding basic research on the mechanisms of DNA mobility in expedient, nonmodel organisms.

REFERENCES

Blackburn EH, Collins K. 2011. Telomerase: An RNP enzyme synthesizes DNA. *Cold Spring Harb Perspect Biol* **3:** 205–213.

Duffy S, Shackelton LA, Holmes EC. 2008. Rates of evolutionary change in viruses: Patterns and determinants. *Nat Rev Genet* **9:** 267–276.

Finnegan DJ. 2012. Retrotransposons. *Curr Biol* **22:** R432–R437.

Lambowitz AM, Zimmerly S. 2004. Mobile group II introns. *Annu Rev Genet* **38:** 1–35.

Lambowitz AM, Zimmerly S. 2011. Group II introns: Mobile ribozymes that invade DNA. *Cold Spring Harb Perspect Biol* **3:** a003616.

Mohr S, Ghanem E, Smith W, Sheeter D, Qin Y, King O, Polioudakis D, Iyer VR, Hunicke-Smith S, Swamy S, et al. 2013. Thermostable group II intron reverse transcriptase fusion proteins and their use in cDNA synthesis and next-generation RNA sequencing. *RNA* **19:** 958–970.

Ng KK, Arnold JJ, Cameron CE. 2008. Structure-function relationships among RNA-dependent RNA polymerases. *Curr Top Microbiol Immunol* **320:** 137–156.