

Reverse transcriptase: from transcriptomics to genome editing

Samara Martín-Alonso, Estrella Frutos-Beltrán and Luis Menéndez-Arias *

Centro de Biología Molecular “Severo Ochoa” (Consejo Superior de Investigaciones Científicas and Universidad Autónoma de Madrid), c/ Nicolás Cabrera 1, Campus de Cantoblanco-UAM, 28049 Madrid, Spain

* Corresponding author E-mail: lmendez@cbm.csic.es

S.M.-A. and E.F.-B. should be considered joint first authors

Samara Martín-Alonso: ORCID no. 0000-0002-5459-7813

Estrella Frutos-Beltrán: ORCID no. 0000-0002-3103-0865

Luis Menéndez-Arias: ORCID no. 0000-0002-1251-6640

Lab website: www.cbm.csic.es/retrovir

Abstract

Reverse transcriptases (RTs) are enzymes with the ability to generate a complementary strand of DNA (cDNA) from RNA. Coupled with a polymerase chain reaction (PCR), RTs have been widely used for the detection of RNAs and for cloning expressed genes. Classical retroviral RTs have been successfully improved by protein engineering. These enzymes and newly characterized RTs are key elements in the development of next-generation sequencing techniques that are now being applied to the study of transcriptomics. In addition, engineered RTs fused to a CRISPR/Cas9 nickase have recently shown great potential as tools to manipulate eukaryotic genomes. In this review, we discuss the properties and uses of wild-type and engineered RTs in biotechnological applications, from conventional RT-PCR to recently introduced prime editing.

Keywords: reverse transcriptase, RNA-seq, fidelity, RT-PCR, prime editors, next-generation sequencing

Technological applications of reverse transcription

Discovered in retroviruses, reverse transcriptases (RTs) convert the viral RNA genome first into complementary DNA (cDNA), and then into double-stranded DNA that integrates into the host genome [1]. Reverse transcription coupled to a polymerase chain reaction (PCR) is a popular technique in molecular biology (known as RT-PCR) and has been widely used for the detection of expressed genes, examination of transcript variants, and generation of cDNA templates for cloning and sequencing (Box 1). One of the first applications of RTs in molecular biology was the construction of cDNA libraries, used in the characterization of novel RNA transcripts, gene sequencing and, after cloning, in the expression of recombinant proteins.

During the last twenty years, a number of techniques using reverse transcription have been developed for large-scale analysis of gene expression. DNA microarrays contain many microscopic spots with DNA bound to a solid surface. These DNAs are usually probes representing sequences of the genome or cDNA of an organism. DNA microarrays are used to determine the expression levels of large numbers of genes simultaneously. For this purpose, total RNA or mRNA from cells or tissues

are labeled with chemiluminiscent or fluorescent dyes, and used as targets for the DNA probes fixed in the microarrays. In experiments involving RNA microarrays, cellular RNAs are reverse transcribed to cDNAs that are then labeled with fluorescent markers before hybridization [2]. Probe-target hybridization is quantified by detection of labeled targets to determine relative abundance of nucleic acid sequences in the target. For these applications, RTs should be able to synthesize long cDNAs even when RNA sequences have a high GC content or secondary structure, and ensure high signal-to-background ratios by efficiently incorporating modified nucleotides during polymerization.

The development of next-generation RNA sequencing (**RNA-seq**) has revolutionized transcriptomics and the analysis of RNA population dynamics. RNA-seq facilitates the analysis of whole transcriptomes of cells, tissues or organisms, including the identification of coding and long non-coding RNAs, evaluation of gene expression levels, discovery of splice variants and fusion transcripts, and the detection of gene expression at a very low level. The achievement of these goals depends on complete and full-length cDNA libraries whose synthesis relies on highly efficient RTs. To maximize efficiency, these enzymes should have desirable properties such as a good catalytic efficiency at high temperatures, ability to synthesize long cDNA products, and high sensitivity (necessary to detect low-abundance transcripts without bias) (Table 1). Other relevant RT features to be considered are their fidelity of DNA synthesis, efficient bypass of DNA damage, and their RNase H, strand transfer and **strand displacement** activities. Additional RT characteristics, such as their ability to incorporate modified nucleotides, and replicate templates with altered backbones might be important for specific applications. In this review we summarize advances in the discovery and characterization of RTs and their use in transcriptomics analysis, synthetic biology and genome editing, while providing an overview of their advantages and limitations while using those technologies.

Currently, most of the protocols aimed to analyze RNA transcripts involve the use of recombinant variants of murine leukemia virus (MLV) RT, although avian myeloblastosis virus (AMV) and human immunodeficiency virus type 1 (HIV-1) RTs, and a thermostable *Geobacillus stearothermophilus* group II intron RT have been extensively characterized as well. The structures of MLV, HIV-1 and *G. stearothermophilus* group II intron RTs are shown in Figure 1, together with the large fragment of Bst DNA polymerase, representing a group of prokaryotic enzymes showing improved isothermal amplification performance (Table 1).

RTs showing good catalytic activity at high temperatures are usually preferred for most biotechnological applications. HIV-1 RT variants show higher activity than wild-type MLV and AMV RTs at 50-60°C [3-5], although other RTs showing activity at high temperatures have been recently described. Examples are MLV RT mutants E69K/E302R/W313F/L435G/N454K [6] and L139P/D200N/T330P/L603W/E607K [7], the HIV-1 group O RT mutant K358R/A359G/S360A [8], and bacterial group II intron RTs [9]. In addition, several companies distribute thermostable RNase H-deficient MLV RTs containing undisclosed combinations of amino acid substitutions (Table 1). Some of those enzymes retain significant activity at temperatures above 75°C [8,10].

Transcriptomics and relevant RT properties

Current technologies to analyze cellular transcriptomes rely on next-generation sequencing (mainly RNA-seq). Progress in next-generation sequencing technologies has paved the way for the advent of single-cell transcriptomics, facilitating the study of the gene expression profile for individual cells under particular biological conditions or stimuli (for example, normal and cancer cells, circulating tumor cells, early events in embryogenesis), leading to a deeper knowledge of cellular heterogeneity [11,12]. RNA-seq methods involve the use of RNA libraries, obtained by following three major steps: (i) RNA isolation from cells, tissues or organisms; (ii) RNA depletion or selection (focused on enrichment in particular RNA fractions, or elimination of unwanted RNA (e.g. ribosomal RNA in most applications); and (iii) cDNA synthesis. Apart from ensuring high-quality reads, representative cDNA libraries should provide a relatively uniform read distribution across transcript length (avoiding over-representation of 5' or 3' gene sequences) and retain the strand specificity of origin for each transcript. Another desirable property of RTs, particularly for single-cell transcriptomics, is their ability to synthesize DNA from minute amounts of RNA.

Fidelity of DNA synthesis

High-fidelity is especially important for single-cell transcriptomics, particularly for the detection and proper characterization of low abundance RNAs, since errors introduced during cDNA synthesis can have a significant effect on the obtained results. However, RTs are DNA polymerases devoid of proofreading activity and show relatively low fidelity in comparison with eukaryotic DNA polymerases [13]. Nevertheless, it is widely accepted that oncoretroviral RTs (MLV or AMV RTs) are 10-15 times more faithful than the prototypic HIV-1 RT (M/B/BH10) and about 6 times more

accurate than HIV-1 group O RT (O/ESP49) [14,15; reviewed in ref. 13]. HIV-1 (M/B/BH10) RT has relatively high error rates, although reported estimates show large differences (6×10^{-5} to 6.7×10^{-4}) [13]. However, these data have been obtained in DNA-dependent DNA polymerization assays (mainly M13 *lacZ* forward mutation assays), carried out at 37°C. Interestingly, in these assays, the HIV-1 O/ESP49 RT showed a lower tendency to introduce **indels** in comparison with AMV, MLV and HIV-1 M/B/BH10 RTs [5,14]. On the other hand, mutants of HIV-1 O/ESP49 RT, such as K65R or K65R/V75I showed 10-fold increased fidelity relative to the WT RT while maintaining its thermal stability [14], while RNase H-deficient mutants of this enzyme (e.g. D443N, E478Q or V75I/E478Q) showed 4.5 to 6.6-fold increased fidelity, although with high frameshift error rates [16].

Unfortunately, the impact of temperature on the accuracy of DNA synthesis (particularly relevant for thermostable RTs) has not been measured. In addition, most of the commercialized RTs contain amino acid substitutions whose impact on fidelity has not been determined. In a comparative study using M13 *lacZ* forward mutation assays, researchers found that the error rate of HIV-1 O/ESP49 RT at 55°C was 1.3×10^{-5} , similar to that shown by a commercial MLV RT (containing mutations E69K/E302R/W313F/L435G/N454K) at its recommended working temperature (i.e. 50°C) [17]. In addition, Mohr and colleagues [10] reported error rates of 0.64 and 0.86×10^{-5} for bacterial group II intron RTs, 2 to 4 times lower than those determined for SuperScript III MLV RT. Those values were obtained from reactions carried out at 55-60°C using as a template a 269-nucleotide in vitro-transcribed RNA encoding a segment of the *lacZ* α -fragment. Biochemical studies carried out with structured and G-rich RNA oligonucleotides revealed that the thermostable *G. stearothermophilus* group II intron RT had a stronger tendency to introduce slippage errors (particularly in homopolymeric runs), in comparison with retroviral RTs including HIV-1 RT [18]. These errors may result in a higher frequency of indel errors.

Interestingly, using M13 *lacZ* forward mutation assays, we have recently shown that RNA-dependent DNA polymerization error rates were similar for all tested RTs, including the most faithful enzymes (e.g. MLV RT, AMV RT or the double-mutant K65R/V75I HIV-1 group O RT) and the error prone HIV-1 M/B/BH10 RT with values in the range of 2.5×10^{-5} to 3.5×10^{-5} [19]. These values are similar to those reported for transcription error rates in different organisms [20-22; reviewed in ref. 23] and suggest that heterogeneities found in the template RNA (i.e. total cellular RNA, mRNAs, etc.) hamper the correct evaluation of cDNA synthesis error rates introduced by RTs. We should keep in mind that errors introduced by T7 RNA polymerase (or equivalent enzymes used

in vitro transcription) contribute to the observed RNA heterogeneity, and this variability affects RT error rate measurements.

Next-generation sequencing combined with the use of primer IDs has allowed overall RT fidelity assessments using different mRNA transcripts as templates. Using primer IDs (i.e., degenerated oligonucleotide tags used as part of specific primers to synthesize the cDNA) to determine consensus sequences, it is possible to identify individual products of the RT reaction and correct errors introduced during PCR resampling and sequencing [24,25]. These methods have been used to determine RT error rates around 10^{-4} for commercial MLV RTs (e.g. SuperScript III and IV) [25,26], as well as wild-type MLV and HIV-1 M/B/BH10 RTs [27,28]. However, up to five-fold differences were observed when comparing data obtained with the same enzyme or in different labs. These variations were attributed to the different template sequences and reaction conditions used in the experiments. Nevertheless, those experiments reveal that higher concentrations of Mg^{2+} (above 1 mM) decrease the accuracy of the RTs [28,29].

Nonetheless, fidelity estimates discussed above made no distinction between transcription and reverse transcription errors and this is an important issue, often overlooked when determining the accuracy of reverse transcription using next-generation sequencing. The introduction of barcodes in the template RNA is necessary to discriminate between errors made during transcription and those introduced while the RT synthesizes the cDNA. One method that takes into account this problem and allows determination of transcription and reverse transcription fidelities has been recently described by Reid-Bayliss & Loeb [30]. After transcription, barcodes are introduced in the RNAs using RNA ligases, and then products are circularized to obtain multiple cDNA copies of each individual mRNA molecule, through rolling-circle reverse transcription. Subsequently, and as described above, each cDNA is tagged again with a second barcode by means of a primer ID, which allows distinguishing cDNA errors from those introduced by the amplification and sequencing steps. This method reveals a transcriptional error rate around 2×10^{-5} [30], although reverse transcription error rates were not reported in these assays.

Template switching and strand specificity

During retroviral reverse transcription, the RT DNA polymerase activity produces both (–) and (+) DNA strands, while the RNase H activity of the enzyme removes the viral genomic (+) strand RNA

template. Two strand transfer events are required to complete this process. **Template switching** plays a major role in the generation of recombinant virus. Interestingly, this RT property has been exploited by different RNA-seq methods to generate full-length cDNA libraries (e.g. STRT [31] or SMART-seq [32-35]). Unlike in reverse transcription where sequence identity is critical for strand transfer, in RNA-seq, template switching is used for the addition of adaptor DNA oligonucleotides containing specific sequences that can be later used for PCR enrichment. This activity eliminates the need for a separate RNA 3'-adaptor ligation step. Library preparation based on template switching has emerged as the method of choice for single cell RNA-seq due to its relatively simple workflow and the low amounts of RNA required [36]. A representative workflow (SMART-seq2 [34]) is shown in Figure 2. In a standard SMART-seq2 protocol, the RT is used to synthesize the first strand cDNA, and anchoring the primer for second strand DNA synthesis which is carried out by a high-fidelity DNA polymerase. However, the RT can also be used to obtain the second strand DNA, as in the case of the Smart-3SEQ method [37]. In these protocols, the higher error rate of the used RT should be taken into account when analyzing the obtained results.

Methods described above take advantage of the ability of MLV RTs to add non-templated C residues to the 3' end of a completed cDNA [32,38]. However, a strong tailing activity has also been recently demonstrated for *G. stearothermophilus* group II intron RT [39] and several HIV-1 group O RT mutants (e.g. K358R/A359G/S360A and RNase H-deficient variants K358R/A359G/S360A/D443N and K358R/A359G/S360A/E478Q) [40], as well as wild-type and engineered MLV RTs [41]. Those polymerases as well as AMV and HIV-1 M/B RTs show a preference for incorporating A at the 3' end of blunt template-primers [38,40,42,43], although extensions are usually limited to one or two nucleotides. Studies carried out with commercial MLV RTs showed that their tailing activity depended on the sequence of the first few nucleotides of the 5' RNA template, and could not be modulated by the sequence of the adaptor (or template switching oligonucleotide) [41].

The 3'-overhangs generated by non-templated addition can interact with 3' ends of RNA or DNA acceptor strands, while RTs stabilize the complex and promote DNA synthesis. The characterization of this **clamp activity** has been reported for several RTs. Thus, HIV-1 M/B RT shows maximum efficiency with two-nucleotide overhangs, containing at least one C or G [44]. However, bovine immunodeficiency virus and the long terminal repeat retrotransposon Tf1 RTs required less rigorous conditions and were prone to template switching even with only single-nucleotide complementarity [45]. Interestingly, *G. stearothermophilus* group II intron RT also showed optimal

template switching in the presence of a single nucleotide 3'-DNA overhang complementary to the 3' nucleotide of the acceptor RNA [39]. Longer overhangs seem to decrease progressively template switching rates for that enzyme. Clamp activities of wild-type MLV and HIV-1 M/B RTs were similar, and in both cases, the phosphorylation of the 5' end of the RNA template produce a significant decrease in template switching [45]. These findings have been recently confirmed and extended for several commercial MLV RTs [41]. However, in these studies, authors also found that a 5'-m⁷G cap in the RNA template was a strong enhancer of the efficiency of template switching, particularly for the sequence m⁷GpppG [41].

Unfortunately, methods used in library construction based on template switching have important drawbacks since uncontrolled strand transfer events lead to the generation of spurious second strand cDNA and antisense artifacts [46,47], and facilitate the synthesis of undesirable recombinants. These recombinants may also originate during cDNA synthesis due to strand transfer events occurring during polymerization. The probability that template switching will occur at a given template position depends on the RT properties, nucleotide concentrations in the reaction, and the presence of secondary structures in the RNA template. Studies carried out with HIV-1 RTs indicate that RNase H inactivation abrogates strand transfer [40]. Therefore, RTs without RNase H activity (e.g. HIV-1 RTs with substitutions D443N or E478Q [40]), or bacterial group II intron RTs [9,10] are expected to increase strand-specificity while preventing the formation of recombinant artifacts, seeing as fusion reads in transcriptomics data.

Strand displacement and gene sequence coverage

In general, template switching-based methods designed for obtaining single-cell transcriptomes (e.g. SMART-seq/SMART-seq2 or Quartz-seq) provide full-length transcript coverage, but are not suitable for the introduction of barcodes (or unique molecular identifiers). Therefore, they are inferior to methods such as STRT-seq, CEL-seq/CEL-seq2, MARSseq, and Drop-seq, that involve the use of barcodes and provide better strand specificity, but introducing 5'- and 3'-biased gene coverages [36]. The strong strand displacement activity shown by retroviral RTs contributes to this bias [48]. Thus, in experiments analyzing human cell mRNAs (from HeLa and MCF-7 cultures), cDNAs obtained with SuperScript III MLV RT at 50°C showed a pronounced 3' bias, while transcript coverage was more evenly distributed in reactions catalyzed by a bacterial group II intron RT at 60°C [10,49]. Although differences between both enzymes were attributed to the higher

processivity of the group II intron RT [10], other factors might also have influenced the results, including the catalytic activity or stability (half-life) of the enzyme in the reaction conditions. These experiments were carried out with oligo(dT) primers, that avoid biases introduced by random hexamer priming in the nucleotide composition at the beginning of transcriptome sequencing [50].

Several methods including the use of linkers that hybridize to the mRNA have been used to circumvent these problems (e.g. Lexogen CORALL and SENSE Total RNA-Seq kits [51-53]). In this context, RTs defective in strand displacement could be useful in future developments of these technologies [54]. On the other hand, the RT strand displacement activity facilitates the introduction of barcodes by taking advantage of rolling circle amplification during cDNA synthesis [30,55,56]. In addition, reverse transcription-rolling circle amplification can be used for the detection of different RNA templates, including circular RNAs [57]. AMV RT and several commercial MLV RTs were found to be suitable for these applications, although ProtoScript II RT seems to be the most efficient for the amplification of circular RNAs [57].

Novel RTs for epitranscriptomics and synthetic biology

Many DNA polymerases have been engineered not only to improve their amplification and sequencing capabilities but also to expand their substrate spectra, particularly by altering their nucleic acid, nucleotide substrate preferences and even base modifications in cellular RNA [recent reviews in refs. 58,59]. A pioneering example was described by Harcourt and colleagues [60] who used *Thermus thermophilus* DNA polymerase I to detect *N*⁶-methyladenosine (m6A) in mRNA. This enzyme acts as an RT in the presence of Mn²⁺ [61], and shows a stronger preference for A versus m6A in comparison with retroviral RTs, in dTTP incorporation assays [60]. A variant of *Thermococcus kodakaraensis* (KOD) exo- DNA polymerase able to discriminate processing of unmodified dGTP between cytosine and 5-methyl-2'-deoxycytosine in the template has also been described [62]. More recently, directed evolution has been used to select an HIV-1 RT variant (containing amino acid substitutions V75F/D76A/F77A/R78K/W229Y/M230L) that displayed high mutation rates at RNA positions containing *N*¹-methyladenosine, thereby facilitating the identification of these RNA modifications in cellular transcriptomes [63].

Increased error rates at RNA positions containing *N*⁶-methyladenosine or 5-hydroxymethyluridine have been reported for reverse transcription reactions catalyzed by retroviral RTs (AMV, MLV and

ProtoScript II RTs). However, other RNA modifications such as 5-methylcytosine, 5-methyluridine or pseudouridine do not seem to affect the cDNA mutational spectra [64]. Recently, Werner and colleagues [65] have shown that the copying efficiency of RNA templates containing modified ribonucleotides (i.e. N^1 -methyladenosine, N^1 -methylguanosine and N^2,N^2 -dimethylguanosine) varied depending on the RT and the modified base detected. Interestingly, their RNA-seq results showed that the most processive enzymes (TGIRT-III, and SuperScript IV and wild-type HIV-1 RTs) were able to produce longer reads when RNAs contained N^1 -methyladenosine, consistently with their lower arrest rates and higher tendency to introduce mismatches opposite to the modified base. The ability of group II intron RTs (e.g. TGIRT-III) to distinctively misincorporate at dimethyl sulfate (DMS)-modified nucleotides without falling off the RNA has been used in DMS-MaPseq, a widely used method for RNA structure mapping [66].

Using an emulsion-based selection method, Ellefson and colleagues [67] obtained a thermostable KOD DNA polymerase with expanded template specificity and a preference for RNA templates. This enzyme showed proofreading activity on both RNA and DNA templates, and increased fidelity in comparison with the wild-type MLV RT (as measured by deep sequencing of the reverse transcription reaction on the *HSPCB* gene) [67]. The evolved KOD DNA polymerase contained 16 amino acid substitutions in comparison with the wild-type enzyme where E664K was critical to increase binding to RNA/DNA heteroduplexes [68]. Exonuclease-deficient KOD DNA polymerase was able to reverse transcribe fully modified templates containing 2'-O-methylated nucleotides, a property also shown by AMV and MLV RTs [69,70].

In general, **xeno nucleic acids** (XNA) such as 2'-O-methylated DNAs, are poor substrates of the DNA polymerase activity of retroviral RTs. However, there are interesting exceptions to this general trend. Thus, wild-type MLV RT and commercialized mutants (e.g. SuperScript II and III RTs) can use locked and α -L-threofuranosyl nucleic acids (LNA and TNA, respectively) as templates of the reverse transcription reaction [71,72], while 2'-5'-isomeric DNA and cyclohexenyl nucleic acids (CeNA) can be substrates of the HIV-1 RT DNA polymerase activity [73,74]. Lelyveld and colleagues [75] have recently shown that $N3' \rightarrow P5'$ phosphoramidate DNA can be used as substrate by several RTs, most notably AMV, MLV and commercially available RNase H-deficient variants ProtoScript II and ThermoScript RTs. However, this seems to be an inefficient and error-prone process. Apart from their potential in synthetic biology, XNAs are interesting in medicinal chemistry for their potential development as aptamers resistant to nucleases. High-affinity XNA aptamers containing 2'-

fluoroarabinonucleotides (FANA) have been successfully selected for binding HIV-1 RT [76]. These molecules were also highly specific and did not bind AMV and MLV RTs.

Prime editors and genome editing technologies

In recent years, **CRISPR/Cas9**-based technologies have revolutionized life sciences by enabling genome editing in many organisms [77]. In this methodology, Cas9 protein introduces a double-strand break in a DNA sequence recognized by a twenty-nucleotide guide sequence, that is subsequently repaired either by the non-homologous end joining (NHEJ) or the homology-directed repair (HDR) systems from the own cell, the last one using as template either the sister chromatid or a provided sequence with the desired change (Box 2). However, achieving any desired change without collateral, undesired effects is an elusive goal [78,79]. In an impressive paper, Anzalone and colleagues [80] described a new technology (known as '**prime editing**') that uses reverse transcription to repair virtually any errors (including all types of base substitutions, as well as small indels) in cell genomes (Figure 3). The key components of the prime editing system are a fusion protein made of a *Streptococcus pyogenes* Cas9 nickase H840A mutant [81] and an engineered MLV RT that incorporates the correct sequence in the target genome, together with a prime editing guide (pegRNA) that serves as template. Around twenty thermostable MLV RTs were used in the development of the system [7], although the best results were obtained with mutant D200N/T306K/W313F/T330P/L603W (Box 2). Using this method, authors were able to introduce more than 175 edits in human cells including targeted deletions, insertions and twelve types of point mutations. By using the Cas9 H840A mutant that contains an inactivated HNH domain, introduction of double-strand breaks in target DNA can be avoided. In general, the method improves efficiency and reduces off-target editing in comparison with base editing methods [82]. However, fundamental questions on the molecular mechanism of the process still remain. For example, we do not know how the RT fidelity could affect the outcome of the experiment. Prime editors contain template pegRNAs that are synthesized by cellular RNA polymerase III, an enzyme that has proofreading activity [83]. Therefore, synthesis of pegRNAs is expected to be more accurate than the synthesis of mRNAs, that depend on RNA polymerase II.

The therapeutic utility of this technique is still uncertain, although we anticipate difficulties in the delivery of the components of the system, while editing efficiency and specificity are yet to be addressed in future studies.

Conclusions and future perspectives

Probably, RTs are among the most relevant enzymes in biotechnology. Further advances in transcriptomics, synthetic genetics and genome editing will require RTs with improved catalytic properties that should increase the sensitivity of RNA detection or the efficiency and specificity of the reverse transcription reaction (see Outstanding questions). RTs with increased **template-primer binding affinity** could be helpful to detect minute amounts of RNA, and facilitate its analysis. However, currently available RTs should still be improved, through the optimization of interactions between the enzyme and the template-primer. Rational design and directed evolution might be useful to achieve this goal. In addition, increased fidelity will be important for improving genome editing, although better methods to measure cDNA synthesis accuracy need to be developed, particularly to set up robust comparisons between different RTs. Measurements of fidelity of DNA synthesis are time consuming and technically difficult due to limitations in our capacity to distinguish between errors found in the RNA used as template and those made by the RT. A fully satisfactory method to determine fidelity of RNA-dependent DNA synthesis is still missing. In the field of synthetic biology, the development of RTs with the capacity of using diverse xeno nucleic acids as templates might be helpful to develop nanocontainers and other molecular devices in nanomedicine. Again, rational design and directed evolution should play a role in these developments. At any rate, our understanding of RT function and activity will be fundamental for further developments in biomedicine, but also to generate biostable ligands and nanostructures with expanded chemical functionality.

Supporting information. Table S1, additional references for characterization and application of enzymes shown on Table 1.

Acknowledgements. This work was supported in part by the Ministry of Science and Innovation of Spain (grant BIO2016-76716-R (AEI/FEDER, UE)), and Consejo Superior de Investigaciones Científicas (2019AEP001). S.M.-A. is a predoctoral fellow of the Ministry of Science and Innovation of Spain (BES-2017-079836). An institutional grant to CBMSO of Fundación Ramón Areces (Madrid, Spain) is also acknowledged.

Conflict of interest. HIV-1 group O RTs mentioned in the manuscript (SunScript™ and Biotools High Retrotranscriptase) and methods for their use are the subject of patents and patent applications that have been licensed by CSIC to 4Basebio AG (previously Sygnis AG/Expedeon Ltd) and Biotools Biotechnological & Medical Laboratories, S.A., respectively. L.M.-A. receives royalty payments for the sale of relevant enzymes and kits.

References:

1. Menéndez-Arias, L. *et al.* (2017) Viral reverse transcriptases. *Virus Res.* 234, 153-176
2. Sealfon, S.C. and Chu, T.T. (2011) RNA and DNA microarrays. *Methods Mol. Biol.* 671, 3-34
3. Gerard, G.F. *et al.* (2002) The role of template-primer in protection of reverse transcriptase from thermal inactivation. *Nucleic Acids Res.* 30, 3118-3129
4. Yasukawa, K. *et al.* (2010) Increase in thermal stability of Moloney murine leukaemia virus reverse transcriptase by site-directed mutagenesis. *J. Biotechnol.* 150, 299-306
5. Álvarez, M. *et al.* (2009) Increased thermostability and fidelity of DNA synthesis of wild-type and mutant HIV-1 group O reverse transcriptases. *J. Mol. Biol.* 392, 872-884
6. Arezi, B. and Hogrefe, H. (2009) Novel mutations in Moloney murine leukemia virus reverse transcriptase increase thermostability through tighter binding to template-primer. *Nucleic Acids Res.* 37, 473-481
7. Baranauskas, A. *et al.* (2012) Generation and characterization of new highly thermostable and processive M-MuLV reverse transcriptase variants. *Protein Eng. Des. Sel.* 25, 657-668
8. Matamoros, T. *et al.* (2013) Major groove binding track residues of the connection subdomain of human immunodeficiency virus type 1 reverse transcriptase enhance cDNA synthesis at high temperatures. *Biochemistry* 52, 9318-9328
9. Stamos, J.L. *et al.* (2017) Structure of a thermostable group II intron reverse transcriptase with template-primer and its functional and evolutionary implications. *Mol. Cell.* 68, 926-939
10. Mohr, 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

11. Shapiro, E. *et al.* (2013) Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat. Rev. Genet.* 14, 618-630
12. Deng, Y. *et al.* (2019) Single-cell omics analyses enabled by microchip technologies. *Annu. Rev. Biomed. Eng.* 21, 365-393
13. Menéndez-Arias, L. (2013) In: Human immunodeficiency virus reverse transcriptase (Le Grice S & Götte M, eds), pp. 225-252. Springer Science+Business Media, New York
14. Barrioluengo, V. *et al.* (2011) Thermostable HIV-1 group O reverse transcriptase variants with the same fidelity as murine leukaemia virus reverse transcriptase. *Biochem. J.* 436, 599-607
15. Barrioluengo, V. *et al.* (2012) Intrinsic DNA synthesis fidelity of xenotropic murine leukemia virus-related virus reverse transcriptase. *FEBS J.* 279, 1433-1444
16. Álvarez, M. *et al.* (2013) Altered error specificity of RNase H-deficient HIV-1 reverse transcriptases during DNA-dependent DNA synthesis. *Nucleic Acids Res.* 41, 4601-4612
17. Álvarez, M. and Menéndez-Arias, L. (2014) Temperature effects on the fidelity of a thermostable HIV-1 reverse transcriptase. *FEBS J.* 281, 342-351
18. Penno, C. *et al.* (2017) Stimulation of reverse transcriptase generated cDNAs with specific indels by template RNA structure: retrotransposon, dNTP balance, RT-reagent usage. *Nucleic Acids Res.* 45, 10143-10155
19. Sebastián-Martín, A. *et al.* (2018) Transcriptional inaccuracy threshold attenuates differences in RNA-dependent DNA synthesis fidelity between retroviral reverse transcriptases. *Sci. Rep.* 8, 627
20. Gout, J.-F. *et al.* (2013) Large-scale detection of in vivo transcription errors. *Proc. Natl. Acad. Sci. U. S. A.* 110, 18584–18589
21. Imashimizu, M. *et al.* (2013) Direct assessment of transcription fidelity by high-resolution RNA sequencing. *Nucleic Acids Res.* 41, 9090–9104
22. Traverse, C.C. and Ochman, H. (2016) Conserved rates and patterns of transcription errors across bacterial growth states and lifestyles. *Proc. Natl. Acad. Sci. U. S. A.* 113, 3311–3316
23. Bradley, C.C. *et al.* (2019) Transcription fidelity: New paradigms in epigenetic inheritance, genome instability and disease. *DNA Repair (Amst.)* 81, 102652

24. Jabara, C.B. *et al.* (2011) Accurate sampling and deep sequencing of the HIV-1 protease gene using a Primer ID. *Proc. Natl. Acad. Sci. U. S. A.* 108, 20166-20171
25. Zhou, S. *et al.* (2015) Primer ID validates template sampling depth and greatly reduces the error rate of next-generation sequencing of HIV-1 genomic RNA populations. *J. Virol.* 89, 8540-8555
26. Zhao, C. *et al.* (2018) An ultraprocessive, accurate reverse transcriptase encoded by a metazoan group II intron. *RNA* 24, 183-195
27. Yasukawa, K. *et al.* (2017) Next-generation sequencing-based analysis of reverse transcriptase fidelity. *Biochem. Biophys. Res. Commun.* 492, 147-153
28. Okano, H. *et al.* (2018) Accurate fidelity analysis of the reverse transcriptase by a modified next-generation sequencing. *Enzyme Microb. Technol.* 115, 81-85
29. Achuthan, V. *et al.* (2014) Human immunodeficiency virus reverse transcriptase displays dramatically higher fidelity under physiological magnesium conditions in vitro. *J. Virol.* 88, 8514-8527. Erratum in: *J. Virol.* (2015) 89, 1944
30. Reid-Bayliss, K.S. and Loeb, L.A. (2017) Accurate RNA consensus sequencing for high-fidelity detection of transcriptional mutagenesis-induced epimutations. *Proc. Natl. Acad. Sci. U. S. A.* 114, 9415-9420
31. Islam, S. *et al.* (2011) Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res.* 21, 1160-1167
32. Zhu, Y.Y. *et al.* (2001) Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction. *Biotechniques* 30, 892-897
33. Ramsköld, D. *et al.* (2012) Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat. Biotechnol.* 30, 777-782. Erratum in: *Nat. Biotechnol.* (2020) 38, 374
34. Picelli, S. *et al.* (2013) Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods.* 10, 1096-1098
35. Bostick, M. *et al.* (2016) Strand-specific transcriptome sequencing using SMART technology. *Curr. Protoc. Mol. Biol.* 116, 4.27.1-4.27.18

36. Ziegenhain, C. *et al.* (2017) Comparative analysis of single-cell RNA sequencing methods. *Mol. Cell.* 65, 631-643.e4
37. Foley, J.W. *et al.* (2019) Gene expression profiling of single cells from archival tissue with laser-capture microdissection and Smart-3SEQ. *Genome Res.* 29, 1816-1825
38. Ohtsubo, Y. *et al.* (2017) Efficient N-tailing of blunt DNA ends by Moloney murine leukemia virus reverse transcriptase. *Sci. Rep.* 7, 41769.
39. Lentzsch, A.M. *et al.* (2019) Template-switching mechanism of a group II intron-encoded reverse transcriptase and its implications for biological function and RNA-Seq. *J. Biol. Chem.* 294, 19764-19784
40. Luczkowiak, J. *et al.* (2018) Template-primer binding affinity and RNase H cleavage specificity contribute to the strand transfer efficiency of HIV-1 reverse transcriptase. *J. Biol. Chem.* 293, 13351-13363
41. Wulf, M.G. *et al.* (2019) Non-templated addition and template switching by Moloney murine leukemia virus (MMLV)-based reverse transcriptases co-occur and compete with each other. *J. Biol. Chem.* 294, 18220-18231
42. Clark, J.M. (1988) Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res.* 16, 9677-9686.
43. Golinelli, M.-P., and Hughes, S. H. (2002) Nontemplate nucleotide addition by HIV-1 reverse transcriptase. *Biochemistry* 41, 5894-5906.
44. Oz-Gleenberg, I. *et al.* (2011) Reverse transcriptases can clamp together nucleic acids strands with two complementary bases at their 3¹-termini for initiating DNA synthesis. *Nucleic Acids Res.* 39, 1042-1053
45. Oz-Gleenberg, I. *et al.* (2012) Template-independent DNA synthesis activity associated with the reverse transcriptase of the long terminal repeat retrotransposon Tf1. *FEBS J.* 279, 142-153
46. van Dijk, E.L. *et al.* (2014) Library preparation methods for next-generation sequencing: tone down the bias. *Exp. Cell Res.* 322, 12-20
47. Levin, J.Z. *et al.* (2010) Comprehensive comparative analysis of strand-specific RNA sequencing methods. *Nat. Methods.* 7, 709-715

48. Vecera, M. *et al.* (2019) Testing of library preparation methods for transcriptome sequencing of real life glioblastoma and brain tissue specimens: A comparative study with special focus on long non-coding RNAs. *PLoS One* 14, e0211978
49. Nottingham, R.M. *et al.* (2016) RNA-seq of human reference RNA samples using a thermostable group II intron reverse transcriptase. *RNA* 22, 597-613
50. Hansen, K.D. *et al.* (2010) Biases in Illumina transcriptome sequencing caused by random hexamer priming. *Nucleic Acids Res.* 38, e131.
51. Yang, S.H. *et al.* (2014) Otx2 and Oct4 drive early enhancer activation during embryonic stem cell transition from naive pluripotency. *Cell Rep.* 7, 1968–1981
52. Lubelsky, Y. and Ulitsky, I. (2018) Sequences enriched in Alu repeats drive nuclear localization of long RNAs in human cells. *Nature* 555, 107-111
53. Ma, F. *et al.* (2019) A comparison between whole transcript and 3' RNA sequencing methods using Kapa and Lexogen library preparation methods. *BMC Genomics* 20, 9
54. Martín-Alonso, S. *et al.* (2020) Defective strand-displacement DNA synthesis due to accumulation of thymidine analogue resistance mutations in HIV-2 reverse transcriptase. *ACS Infect. Dis.* 6, 1140-1153
55. Lou, D.I. *et al.* (2013) High-throughput DNA sequencing errors are reduced by orders of magnitude using circle sequencing. *Proc. Natl. Acad. Sci. U. S. A.* 110, 19872-19877
56. Acevedo, A. and Andino, R. (2014) Library preparation for highly accurate population sequencing of RNA viruses. *Nat. Protoc.* 9, 1760-1769
57. Liu, Y. *et al.* (2020) Direct detection of circRNA in real samples using reverse transcription-rolling circle amplification. *Anal. Chim. Acta.* 1101, 169-175
58. Coulther, T.A. *et al.* (2019) Engineering polymerases for new functions. *Trends Biotechnol.* 37, 1091-1103
59. Taylor, A.I. *et al.* (2019) Beyond DNA and RNA: The expanding toolbox of synthetic genetics. *Cold Spring Harb. Perspect. Biol.* 11, a032490
60. Harcourt, E.M. *et al.* (2013) Identification of a selective polymerase enables detection of N(6)-methyladenosine in RNA. *J. Am. Chem. Soc.* 135, 19079-19082

61. Myers, T.W. and Gelfand, D.H. (1991) Reverse transcription and DNA amplification by a *Thermus thermophilus* DNA polymerase. *Biochemistry* 30, 7661-7666
62. Huber, C. *et al.* (2016) 5-methylcytosine-sensitive variants of *Thermococcus kodakaraensis* DNA polymerase. *Nucleic Acids Res.* 44, 9881-9890
63. Zhou, H. *et al.* (2019) Evolution of a reverse transcriptase to map N(1)-methyladenosine in human messenger RNA. *Nat. Methods* 16, 1281-1288
64. Potapov, V. *et al.* (2018) Base modifications affecting RNA polymerase and reverse transcriptase fidelity. *Nucleic Acids Res.* 46, 5753-5763
65. Werner, S. *et al.* (2020) Machine learning of reverse transcription signatures of variegated polymerases allows mapping and discrimination of methylated purines in limited transcriptomes. *Nucleic Acids Res.* 48, 3734-3746
66. Zubradt, M. *et al.* (2017) DMS-MaPseq for genome-wide or targeted RNA structure probing in vivo. *Nat. Methods* 14, 75-82
67. Ellefson, J.W. *et al.* (2016) Synthetic evolutionary origin of a proofreading reverse transcriptase. *Science* 352, 1590-1593
68. Cozens, C. *et al.* (2012) A short adaptive path from DNA to RNA polymerases. *Proc. Natl. Acad. Sci. U. S. A.* 109, 8067-8072
69. Pan, T. (1995) Novel RNA substrates for the ribozyme from *Bacillus subtilis* ribonuclease P identified by in vitro selection. *Biochemistry* 34, 8458-8464
70. Burmeister, P.E. *et al.* (2005) Direct in vitro selection of a 2'-O-methyl aptamer to VEGF. *Chem. Biol.* 12, 25-33
71. Chaput, J.C. *et al.* (2003) DNA polymerase-mediated DNA synthesis on a TNA template. *J. Am. Chem. Soc.* 125, 856-857
72. Crouzier, L. *et al.* (2012) Efficient reverse transcription using locked nucleic acid nucleotides towards the evolution of nuclease resistant RNA aptamers. *PLoS One* 7, e35990
73. Sinha, S. *et al.* (2004) 2',5'-linked DNA is a template for polymerase-directed DNA synthesis. *J. Am. Chem. Soc.* 126, 40-41

74. Kempeneers, V. *et al.* (2005) Investigation of the DNA-dependent cyclohexenyl nucleic acid polymerization and the cyclohexenyl nucleic acid-dependent DNA polymerization. *Nucleic Acids Res.* 33, 3828-3836
75. Lelyveld, V.S. *et al.* (2019) DNA polymerase activity on synthetic N3'→P5' phosphoramidate DNA templates. *Nucleic Acids Res.* 47, 8941-8949
76. Alves Ferreira-Bravo, I. *et al.* (2015) Selection of 2'-deoxy-2'-fluoroarabinonucleotide (FANA) aptamers that bind HIV-1 reverse transcriptase with picomolar affinity. *Nucleic Acids Res.* 43, 9587-9599
77. Hsu, P.D. *et al.* (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157, 1262-1278
78. Fu, Y. *et al.* (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* 31, 822–826
79. Tsai, S.Q. *et al.* (2015) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* 33, 187–197
80. Anzalone, A.V. *et al.* (2019) Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 576, 149-157
81. Jiang, F. and Doudna, J.A. (2017) CRISPR-Cas9 structures and mechanisms. *Annu. Rev. Biophys.* 46, 505-529
82. Gaudelli, N.M. *et al.* (2017) Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* 551, 464-471
83. Alic, N. *et al.* (2007) Selectivity and proofreading both contribute significantly to the fidelity of RNA polymerase III transcription. *Proc. Natl. Acad. Sci. U. S. A.* 104, 10400-10405
84. Wong, Y.P. *et al.* (2018) Loop-mediated isothermal amplification (LAMP): a versatile technique for detection of micro-organisms. *J. Appl. Microbiol.* 124, 626-643
85. Bower, N.I. and Johnston, I.A. (2010) Targeted rapid amplification of cDNA ends (T-RACE)—an improved RACE reaction through degradation of non-target sequences. *Nucleic Acids Res.* 38, e194-e194

86. Adkar-Purushothama, C.R. *et al.* (2017) 3' RNA ligase mediated rapid amplification of cDNA ends for validating viroid induced cleavage at the 3' extremity of the host mRNA. *J. Virol. Methods* 250, 29-33
87. Jinek, M. *et al.* (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816-821
88. Zhang, F. *et al.* (2012) Engineering of systems, methods and optimized guide compositions for sequence manipulation. U.S. Provisional Patent Application 61/736,527, filed December 12, 2012; later published as WO2014093712A9
89. Gasiunas, G. *et al.* (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 109, E2579-E2586
90. Cong, L. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819-823
91. Komor, A.C. *et al.* (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420-424
92. Rees, H.A. and Liu, D.R. (2018) Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Genet.* 19, 770-788
93. Molla, K.A. and Yang, Y. (2019) CRISPR/Cas-Mediated Base Editing: Technical Considerations and Practical Applications. *Trends Biotechnol.* 37, 1121-1142
94. Doman, J.L. *et al.* (2020) Evaluation and minimization of Cas9-independent off-target DNA editing by cytosine base editors. *Nat. Biotechnol.* 38, 620-628

Box 1. RT-PCR techniques

RT-PCR reactions can be completed in one or two steps, using oligo(dT), random primers or gene-specific primers, and are suitable for quantitative analysis of gene abundance (i.e. RT-qPCR). Quantitative PCR (qPCR) uses real-time fluorescence to measure the quantity of DNA present at each cycle during a PCR, and can be modified to detect and quantify RNA by adding a reverse transcriptase step before the qPCR assay. RT-LAMP (reverse transcription loop-mediated isothermal amplification) is a technique similar to RT-PCR, but it is performed at a constant temperature

between 60 and 65°C, using a set of gene-specific primers and a DNA polymerase with strong strand displacement activity [84].

The identification of 5' and 3' untranslated regions, including transcriptional start sites and promoter regions, has been possible through the development of PCR-based methods such as the rapid amplification of cDNA ends (RACE). In RACE, mRNAs are reverse transcribed to DNA using either a gene-specific primer (antisense oligonucleotide primer) that recognizes a known sequence (5' RACE and 3' RACE) and an oligo(dT) primer annealed to the poly(A) tail of the mRNA, in the case of 3' RACE. In 5' RACE, the 3' end is extended by the RT that adds a homopolymeric tail (usually containing cytosines), or is ligated to an oligonucleotide adaptor primer. Other RACE protocols such as the targeted rapid amplification of cDNA ends (T-RACE), use dUTP-containing PCR primers and uracil-DNA glycosylase to obtain more specific amplifications of 3' and 5' ends of cDNAs, in a single reaction [85]. The RNA ligase-mediated rapid amplification of cDNA ends (5' and 3' RML-RACE) has been developed for the validation of direct cleavage of a target gene by a microRNA (miRNA), and to identify RNA ends when cleavage takes place in the 3' extremity of the target RNA [86]. In this method an RNA adaptor is ligated to the 3' end of the cleaved RNA. The adaptor is adenylated at the 5' end and has a 3' amine group that prevents self-ligation of the oligonucleotide.

Box 2. Design and development of prime editors

Genome editing based on CRISPR-Cas technology was introduced in 2012, and immediately recognized as a revolutionary breakthrough in biomedicine [77,87-90]. CRISPR/Cas9 protein – RNA complexes target a DNA sequence through base pairing with a guide RNA, and create a double-strand break at the locus specified by the guide RNA. Then, non-homologous end-joining (NHEJ) or homology-directed repair (HDR) systems in the cell facilitate genome editing. CRISPR-Cas9 gave researchers the ability to generate targeted random gene disruption, but precision of genome editing has been a question of great concern. Improved accuracy was achieved by conjugating the Cas nuclease with a nuclease deaminase enzyme, and in some cases, a DNA glycosylase inhibitor. These 'base editors' can be successfully used to introduce specific and precise point mutations at a target locus in RNA or DNA [91-94]. However, base editors can make only a subset of edits, while showing preference for some specific mutations, while many sequences are resilient to modification.

These limitations have been overcome by prime editors. For the optimization of these editing tools, Anzalone and colleagues [80] successively used three types of prime editors. The prototype was PE1 and included a Cas9 H840 nickase fused to wild-type MLV RT that together with pegRNA was already successful in introducing transversions in the target sequence. The mutant Cas9 nickase activity acts only on the strand to be edited by the RT, leaving the complementary chain unchanged. PE2 represented a further improvement in the fusion protein and contained a mutant MLV RT (D200N/T306K/W313F/T330P/L603W). Compared to PE1, PE2 showed increased editing efficiency at tested genomic loci. In addition to the PE2 fusion protein and the corresponding pegRNA, PE3 systems use a single guide RNA (sgRNA). The sgRNA can be specifically directed towards a sequence located of the non-edited 14-116 nucleotides away from the site of the pegRNA-induced nick (PE3) or directly to the newly-edited sequence (PE3b), where cellular mismatch repair pathways can fix the information introduced in the edited strand. Compared to PE2, the PE3b strategy showed 3-fold increased editing efficiency and lower levels of indel formation (Figure 3).

Glossary

Clamp activity: is the property of RTs that can simultaneously bind a template-primer hybrid with a 3'-DNA overhang of one to three oligonucleotides, and an acceptor template with a 5' end complementary to the 3' end of the DNA primer. This property allows RT to switch templates while extending the protruding DNA primer.

CRISPR/Cas9: clustered regularly interspaced short palindromic repeats (CRISPR) are DNA sequences found in the genomes of prokaryotes that are used to detect and eliminate foreign DNA. Cas9 is a CRISPR-associated nuclease that recognizes and cleaves specific strands of DNA that are complementary to the CRISPR sequence. Cas9 contains two catalytic domains (RuvC and HNH), and each one cleaves one DNA strand. CRISPR/Cas9 technologies use a predesigned single-guide RNA (sgRNA), containing a specific RNA sequence that recognizes the target DNA region of interest and directs the Cas9 nuclease there for editing.

Indel: Insertion or deletion of nucleotides detected in the genome of an organism, or by extension in assays screening for mutations introduced by RTs and other polymerases.

Prime editing: ‘search-and-replace’ genome editing technology that allows genome rewriting with almost any intended change. It uses a mutant Cas9 nickase (*Streptococcus pyogenes* Cas9 H840A) fused to an RT (wild-type or engineered MLV RT) and a prime editing guide RNA (pegRNA) that encodes the desired genetic information.

Processivity: a property that describes the ability of enzymes to catalyze consecutive reactions without releasing its substrate. For RT and DNA polymerases in general, it is defined as the average number of nucleotides added to the synthesized DNA in a single polymerase binding event.

RNA-seq: techniques that use next-generation sequencing to detect and quantify RNA found in biological samples at a specific moment, facilitating the analysis of cellular transcriptomes.

Strand displacement: a property that describes the ability of RT and other DNA polymerases to displace downstream DNA encountered during templated DNA synthesis.

Template-primer binding affinity: measures the strength of the interaction between template-primer and the RT (or another polymerase). Usually measured and reported by the equilibrium dissociation constant (K_D). Smaller K_D are associated with greater binding affinities.

Template switching: is a mechanism by which RTs are able to change templates while synthesizing DNA. During retroviral reverse transcription this process occurs at least twice (i.e. strand transfer events) and may eventually produce recombinant DNAs.

Xeno nucleic acids: synthetic nucleic acid analogues that have a different sugar backbone than the natural nucleic acids DNA and RNA. Fluoro arabino nucleic acid (FANA), α -L-threofuranosyl nucleic acid (TNA), 1,5-anhydrohexitol nucleic acid (HNA), glycol nucleic acid (GNA), cyclohexenyl nucleic acid (CeNA), or locked nucleic acid (LNA) are examples of different sugar backbones found in xeno nucleic acids. LNAs contained modified RNA nucleotides in which their ribose moieties are modified with an extra bridge connecting the 2' oxygen and 4' carbon.

Highlights

Robust catalytic activity (even at high temperatures) and high fidelity and template-primer affinity are desirable reverse transcriptase (RT) properties in most biotechnological applications.

Engineered murine leukemia virus RTs are most commonly used enzymes, although enzymes from other retroviruses (AMV or HIV-1) and bacterial group II intron RTs are also efficient in most applications.

New technologies using reverse transcription and showing promise in life science research and technology are RNA-seq (transcriptomics analysis), epitranscriptomics and synthetic biology, and genome editing (through the use of prime editors).

Outstanding questions

Available RTs have limitations concerning their ability to synthesize cDNA in the presence of small amounts of RNA and better enzymes will be needed if we wish to obtain accurate and reliable single cell transcriptomes.

The fidelity of RTs has not been correctly evaluated in RNA-dependent DNA polymerase reactions, and this is particularly relevant if we wish to obtain a highly accurate RT. In addition, for most enzymes, fidelity assessments have not been obtained in the same conditions and at the appropriate working temperature.

In general, xeno nucleic acids are poor templates of RTs. How can they best be developed for synthetic biology applications?

Prime editing characterization is still limited and mechanistic details are not fully understood. Still we do not know what are the rules of prime editing, where errors are more likely to occur and the impact of the RT in the generation of indels and off-target edits.

LEGENDS TO FIGURES

Figure 1. Structures of representative RTs. Conserved RT subdomains are similarly colored in each structure. **(A)** MLV RT is an 80-kDa monomer (coordinates taken from Protein Data Bank (PDB) ID: 5DMQ.PDB, including residues 24 to 663). **(B)** HIV-1 RT is a heterodimeric enzyme composed of subunits of 66 and 51 kDa (p66 and p51, respectively) (PDB ID: 1RTD.PDB). The catalytic p66 subunit (residues 1-554) is colored to indicate its different domains and subdomains. The 51-kDa subunit (residues 4-429) is shown using a lightly colored cartoon representation. **(C)** Thermostable *G. stearothermophilus* group II intron RT (PDB ID: 6AR1.PDB) (includes coordinates for residues 2 to 418). RNase H and connection domains are not present in this RT which,

in addition lacks the additional C-terminal DNA endonuclease domain present in most subgroup IIA and IIB intron RTs. **(D)** *G. (Bacillus) stearothermophilus* DNA polymerase I large fragment (PDB ID: 1XWL.PDB; residues 297 to 876). Molecules of nucleic acids and other protein structures present in the corresponding PDB files are not shown for clarity. Molecular graphics images were produced using the UCSF Chimera package.

Figure 2. Outline of the SMART-Seq2 workflow. The SMART-Seq technology takes advantage of the tailing, clamp and template switching activity of RTs. The different steps used in library generation are indicated. Minus and plus strand DNA synthesis is carried out using cDNA synthesis primers and template switching oligonucleotides (TSO), respectively. The cDNA synthesis primers are composed of TTTT-Adaptor1-ISPCR sequences, where ISPCR (standing for “in situ PCR”) is a primer used for amplification following the retrotranscription reaction. TSOs contain a sequence made of ISPCR-Adaptor2-rGrG+G, where the last G is a modified guanosine to produce a locked nucleic acid (LNA). Adaptors are barcodes annealed to oligo(dT) or rGrG+G primers. Tagmentation involves Tn5 transposase cleavage and tagging of the double-stranded DNA, followed by a gap-filling step during PCR amplification to obtain a final indexed DNA. After tagmentation, resulting DNA fragments are flanked by platform-specific sequences (PS1, PS2), indexes (i1, i2) and sequencing primer sites (SP1, SP2).

Figure 3. Prime editors and prime editing mechanism. Prime editors are composed of a Cas9 H840A nickase from *S. pyogenes*, fused to WT MLV RT (in prime editor 1) or engineered MLV RT (prime editors 2, 3 and 3b), and pegRNA (prime editing guide RNA). After hybridization of the target DNA with the pegRNA protospacer sequence, Cas9 H840A makes a single-strand break, 3 nucleotides upstream of the protospacer adjacent motif (PAM). Cas9 recognizes the PAM sequence 5'-NGG-3' (where “N” can be any nucleotide base). After primer binding site hybridization and extension of the DNA by the MLV RT, prime editors 1 and 2 generate an edit-containing 3' flap that can be converted to a 5' flap through flap equilibration. The excision of the 5' flap, followed by ligation and replication fix the edits into the target DNA. Prime editors 3 and 3b enhance editing efficiency by including an additional sgRNA (single guide RNA) that facilitates nicking of the unedited strand, thereby triggering the mismatch repair machinery for fixation of the intended edit. Editing efficiency can be improved by targeting the additional sgRNA to the edited sequence or close to it (prime editor 3b).

Table 1. Characteristics and applications of commercially available RTs

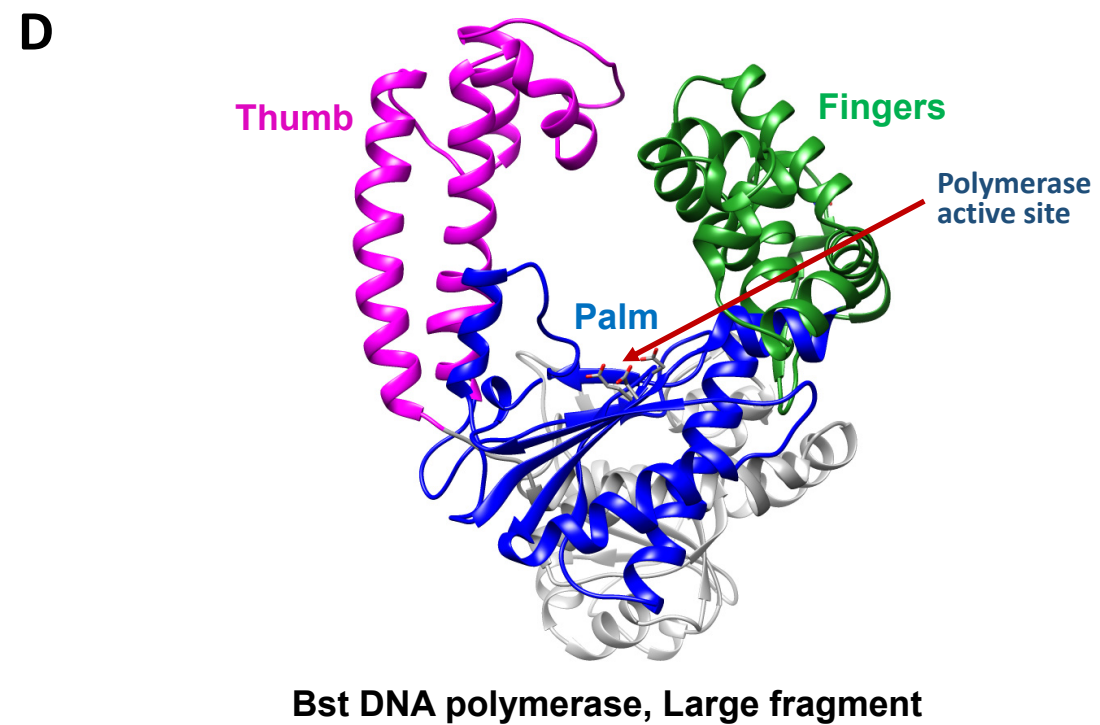
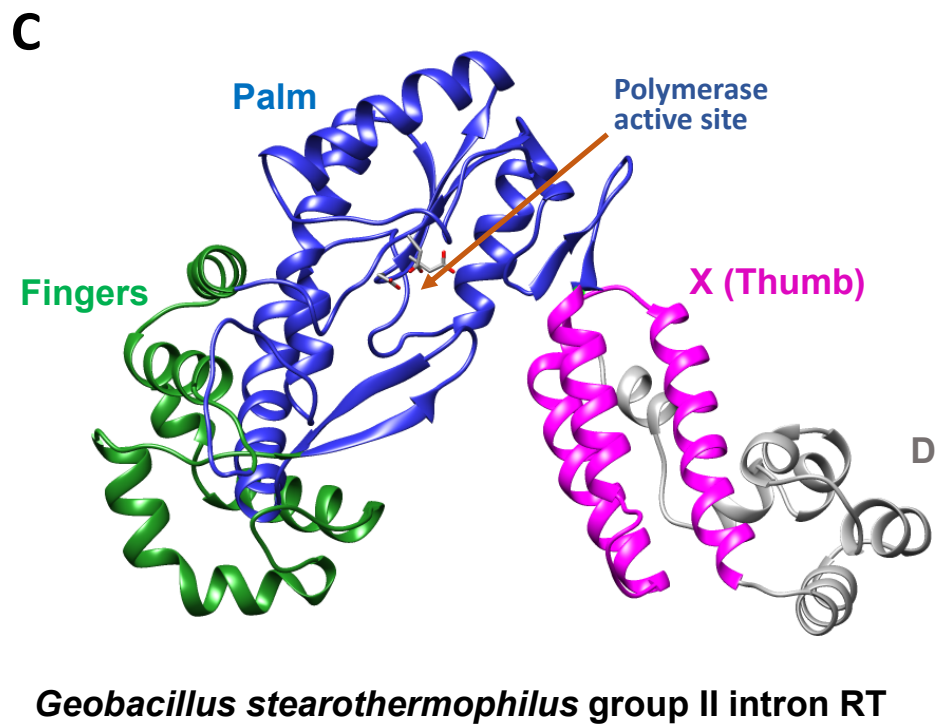
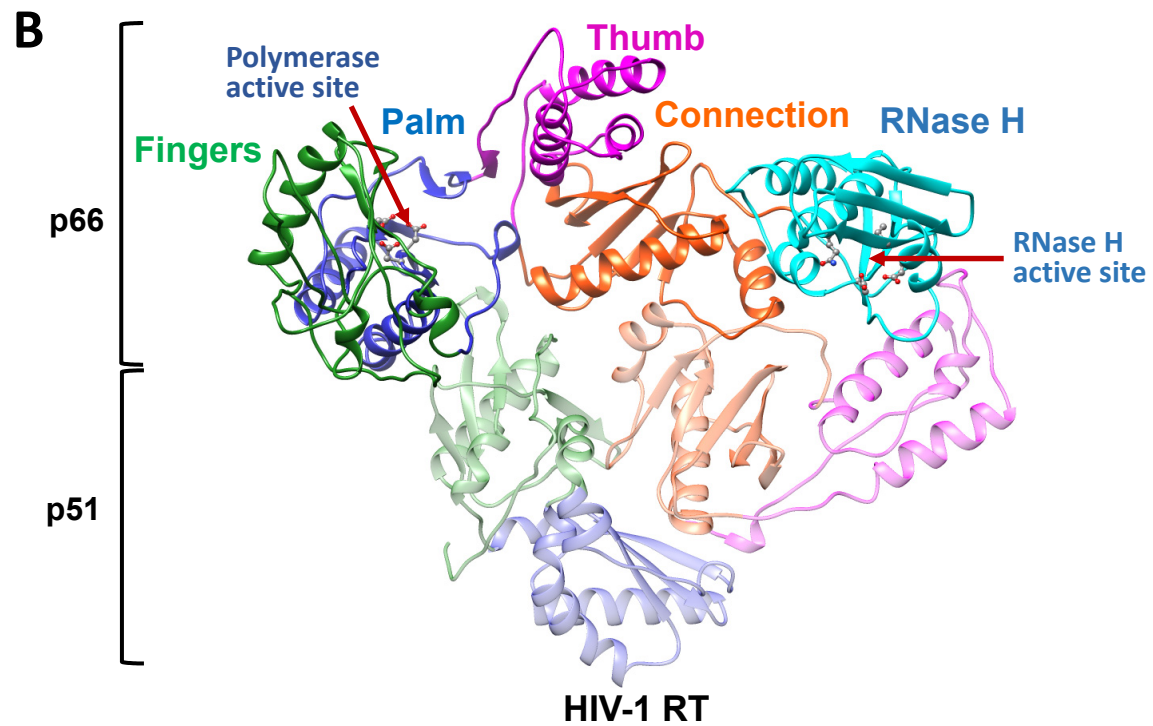
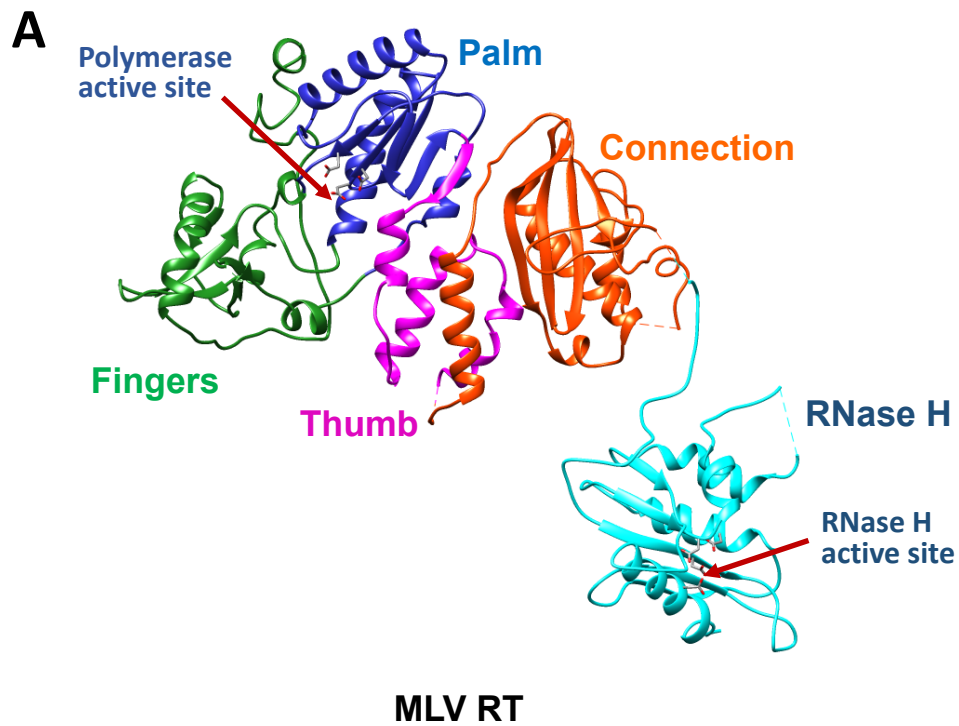
Enzymes	Commercial name (Company)	Structure and RNase H activity	Reaction temperatures (°C)		Amplicon size (kb)	RNA detection levels		Applications
			Optimal	Maximum		Total	mRNA	
AMV RTs	AMV Reverse Transcriptase (New England Biolabs, Promega, Clontech-Takara)	Wild-type RNase H plus	37–58	60	-	1 ng – 1 µg	50 pg – 100 ng	Epitranscriptomics
	eAMV™ (Sigma-Merck)	Engineered RNase H plus	42–65	65	< 14.1	<10 ng/µl	–	SC RNA-seq SC transcriptomics
	ThermoScript™ (Invitrogen– Thermo Fisher Scientific) *	Engineered Reduced RNase H activity	50–65	–	< 12	10 pg – 5 µg	10 pg – 500 ng	Epitranscriptomics Quantification of miRNAs
MLV RTs	M-MLV Reverse transcriptase (Invitrogen-Thermo Fisher Scientific, New England Biolabs, Promega, Sigma-Merck)	Wild-type RNase H plus	37	42	< 7	1 ng – 5 µg	1 ng –500 ng	SC RNA-seq SC transcriptomics Epitranscriptomics
	GoScript™ (Promega)	Wild-type RNase H plus	37–55	–	< 8.9	>5 µg	>500 ng	RNA-seq SC RNA-seq Quantification of miRNAs and lncRNAs
	MultiScribe™ (Thermo Fisher Scientific)	Wild-type RNase H reduced	37	–	< 7	1 ng – 5 µg	1 ng –500 ng	RT-PCR
	AccuScript Hi-Fi (Stratagene-Agilent)	Engineered RNase H minus (Includes 3'-5' exonuclease activity)	37–42	–	< 20	10 ng – 5000 ng	0.1 ng – 100 ng	SC RNA-seq SC transcriptomics
	Affinity Script (Agilent)	Engineered mutant E69K/E302R/W313F/L435G/N454K Unspecified RNase H activity	37–55	55	< 20	1 ng – 500 ng	1 ng – 250 ng	SC RNA-seq
	ArrayScript™ (Invitrogen-Ambion-Thermo Fisher Scientific)	Engineered Unspecified RNase H activity	42	–	< 9	1 µg – 5 µg	10 ng – 100 ng	Microarray SC RNA-seq
	BioScript™ (Bioline)	Engineered Reduced RNase H activity	37–42	45	–	> 5 µg	> 0.5 µg	Library construction
	CycleScript™ (Bioneer)	Recombinant (not specified)	37–55	–	< 9	1 µg	5 ng – 100 ng	Library construction
	EnzScript™ (Biozym)	Engineered RNase H minus	37–55	55	< 12	< 1 pg	–	SC RNA-seq
	EpiScript™ (Lucigen-Biosearch Technologies)	Engineered RNase H minus	37–42	55	< 15	50 pg	–	Epitranscriptomics
	Expand™ Reverse Transcriptase (Roche-Merck)	Engineered RNase H reduced	43	–	<13.5	1 µg	50 ng – 100 ng	RT-PCR
	FIREScript (Solis Biodyne)	Engineered RNase H plus	37	60	–	0.1 ng – 5 µg	–	cDNA synthesis
	GrandScript (TATAA Biocenter) *	Engineered RNase H plus	42	–	–	1 µg – 10 pg	–	SC transcriptomics
	iScript™ (BioRad) *	Engineered RNase H plus	46	–	< 1	100 fg – 1 µg	–	SC RNA-seq SC transcriptomics
	Maxima™ RT (ThermoFisherScientific)	Engineered RNase H plus & minus	50–55	65	< 20	1 pg – 5 µg	0.1 pg – 500 ng	SC RNA-seq SC transcriptomics
	MonsterScript™ (Epicentre-Ecogen)	Engineered RNase H minus	50–60	65	≥ 15	1 pg	–	Library construction RNA-seq
	PrimeScript™ (Takara)	Engineered RNase H minus	42	50	< 12	≤ 5 µg	≤ 1 µg	SC RNA-seq SC transcriptomics
	PrimeScript™ II (Takara)	Engineered RNase H minus	42	50	< 13	≤ 5 µg	≤ 1 µg	SC RNA-seq
	PrimeScript™ III (Takara) *	Engineered RNase H minus	42	55	-	0.1 pg – 1 µg	–	SC RNA-seq
	PrimeScript™ IV (Takara) *	Engineered RNase H minus	42	55	< 13	≤ 5 µg	≤ 1 µg	RT-PCR

	ProtoScript® (New England Biolabs) *	Recombinant (not specified) RNase H plus	42	45	< 10	1 ng – 1 µg	0.1 ng – 100 ng	RNA-seq
	ProtoScript® II (New England Biolabs)	Engineered RNase H reduced	42	48	< 12	1 ng – 1 µg	1 ng – 100 ng	circRNA amplification RNAseq SC RNA-seq
	qScript (Quantabio) *	Engineered RNase H plus	42–52	54	< 1	10 pg – 1 µg	–	SC RNA-seq SC transcriptomics
	RevertAid™ (Thermo Fisher Scientific)	Engineered (>15 amino acid changes) RNase H plus & minus	42–50	55	< 13	0.1 ng – 5 µg	10 pg – 500 ng	RNA-seq SC RNA-seq
	ReverTra Ace® (Toyobo)	Engineered RNase H minus	42–55	55	≥ 14	0.1 µg – 1 µg	50 ng – 500 ng	RNA-seq
	RevertUp II™ (Biotechrabbit)	Engineered RNase H minus	42–55	60	≥ 19	0.1 µg – 1 µg	50 ng – 500 ng	RNA-seq
	Rocketscript™ (Bioneer)	Engineered RNase H plus & minus	42–70	70	> 10	1 pg – 1 µg	–	Library construction
	Script (Jena Bioscience)	Engineered RNase H minus	42–55	–	< 10	10 pg – 5 µg	10 pg – 500 ng	RT-PCR
	SMART® (Clontech-Takara)	Recombinant (not specified)	42	–	< 11.7	5 ng – 5 µg	10 ng – 1 µg	RT-qPCR Microarrays
	SMARTScribe™ (Clontech-Takara)	Engineered Unspecified RNase H activity	42	–	< 14.7	5 ng – 2 µg	10 ng – 1 µg	RNA-seq SC RNA-seq SC transcriptomics Microarrays
	SuperScript™ II (Invitrogen-Thermo Fisher Scientific)	Engineered D524G/D583N/E562Q RNase H reduced	37–42	50	≤ 12.3	1 ng– 5 µg	1 ng– 500 ng	Epitranscriptomics SC RNA-seq SC transcriptomics Microarrays
	SuperScript™ III (Invitrogen-Thermo Fisher Scientific)	Engineered H204R/V223H/T306K/ F309N/D524G/D583N /E562Q RNase H reduced	50	55	> 12	10 pg – 5 µg	10 pg – 500 ng	RNA-seq Microarrays SC transcriptomics
	SuperScript™ IV (Invitrogen-Thermo Fisher Scientific)	Engineered (>15 amino acid changes) RNase H reduced	50–55	65	≤ 12.3	10 pg – 5 µg	10 pg – 500 ng	SC transcriptomics Epitranscriptomics
	Transcriptor Reverse Transcriptase (Roche-Merck)	Engineered RNase H plus	42–65	65	< 14	10 ng – 5 µg	1 ng – 100 ng	Microarrays miRNA expression
HIV-1 RTs	HIV-1 RT (Abcam, Ambion-Thermo Fisher Scientific, Worthington)	Wild-type of group M subtype B RNase H plus	37–42	50	–	–	–	Epitranscriptomics RNA-seq
	Bitools High Retrotranscriptase (Bitools)	Engineered group O variant (K65R/V75I) RNase H plus	45–47	65	1.2	–	0.02 ng	RT-qPCR Synthetic biology
	Sunscript® (4Basebio)	Engineered group O variants with changes K358R/A359G/S360A RNase H plus & minus	65	85	11	–	100 pg	RT-qPCR
Bacterial group II intron RTs	Marathon RT (Kerabast)	Wild-type (<i>Eubacterium rectale</i>) Lacks RNase H domain	42	47	< 10	–	200 ng – 300 ng	RNA-seq
	TGIRT®-III RT (InGex)	Wild-type (<i>Geobacillus stearothermophilus</i>) Lacks RNase H domain	42–70	–	≥ 5 kb	1 ng – 50 ng	–	Epitranscriptomics RNA-seq
Bacterial DNA polymerases	BcaBEST Polymerase (Takara) *	Engineered (<i>Bacillus caldotenax</i> DNA polymerase without 5'-3' and 3'-5' exonuclease activity) Lacks RNase H domain	65	–	< 5	500 ng	–	RT-PCR

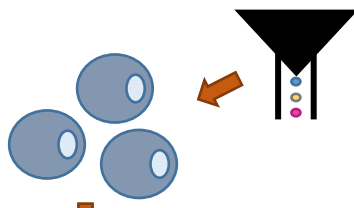
	Bst 3.0 DNA Polymerase (New England Biolabs)	<i>Geobacillus stearothermophilus</i> DNA polymerase I, large fragment (lacks 5'-3' and 3'-5' exonuclease activity) Lacks RNase H domain	65–72	75	≥ 0.3	> 10 copies [^]	–	RT-LAMP
	RapiDxFire™ Reverse Transcriptase (Lucigen-Biosearch Technologies)	Obtained from hot spring bacteria Lacks RNase H domain	55–80	90	≤ 1	≤100 copies [^]	–	cDNA synthesis RT-LAMP
	Volcano2G DNA polymerase (myPOLS Biotec)	Engineered <i>Thermus aquaticus</i> DNA polymerase Lacks RNase H domain	50–95	95	< 0.3	1 ng – 1000 ng	–	RT-qPCR
	Volcano3G DNA polymerase (myPOLS Biotec) *	Engineered <i>Thermus aquaticus</i> DNA polymerase Lacks RNase H domain	58–70	95	< 0.3	0.1 – 2500 ng	–	RT-qPCR
Source not available	SOLIScript (Solis Biodyne)	Engineered RNase H reduced	50	60	–	0.1 ng – 1 µg	–	cDNA synthesis
	Omniscrypt® (Qiagen) *	Heterodimeric RT RNase H plus	37–42	65	< 1.2	50 ng – 2 µg	–	RT-qPCR RNA-seq
	SensiScript® (Qiagen) *	Heterodimeric RT RNase H plus	37–42	65	–	< 50 ng	–	SC transcriptomics

References for each enzyme are provided in the supplementary Table S1. All RTs listed are commonly used for cDNA synthesis and RT-PCR assays. The application list in the last column provides other relevant uses of these enzymes, as found in the scientific literature and in materials provided by their suppliers. In addition to the listed RTs, there are others that have not been characterized so extensively, but are included in different types of kits, usually designed for transcriptomics analysis: 3B High Retrotranscriptase (Indiamart), HotStaRT-Script reverse (Qiagen), ImProm™-II (Promega), ipsogen® RT (Qiagen), Isotherm2G DNA polymerase (myPOLS Biotec), LunaScript™ (New England Biolabs), miScript RT (Qiagen), NEXTFLEX® RNA-Seq Rapid Reverse transcriptase (Perkin Elmer), NEXTFLEX® RNA-Seq Rapid Reverse transcriptase 2.0 (Perkin Elmer), QuantiNova Reverse transcriptase (Qiagen), Quantiscrypt Reverse transcriptase (Qiagen), RetroAMPTM RT DNA Polymerase (Ecogen), SEQzyme (BioRad), T-Script Enzyme (Qiagen), and WarmStart RTx® Reverse transcriptase (New England Biolabs).

Asterisks indicate that the RT is not available as a stand-alone enzyme format, being the simplest packaging a Supermix or a kit. [^], refers to copies of a reference RNA; –, not available. Abbreviations: circRNA, circular RNA; lncRNA, long non-coding RNA; miRNA, micro RNA; SC, single cell.



1. Cells separated by fluorescence-activated cell sorting (FACS)



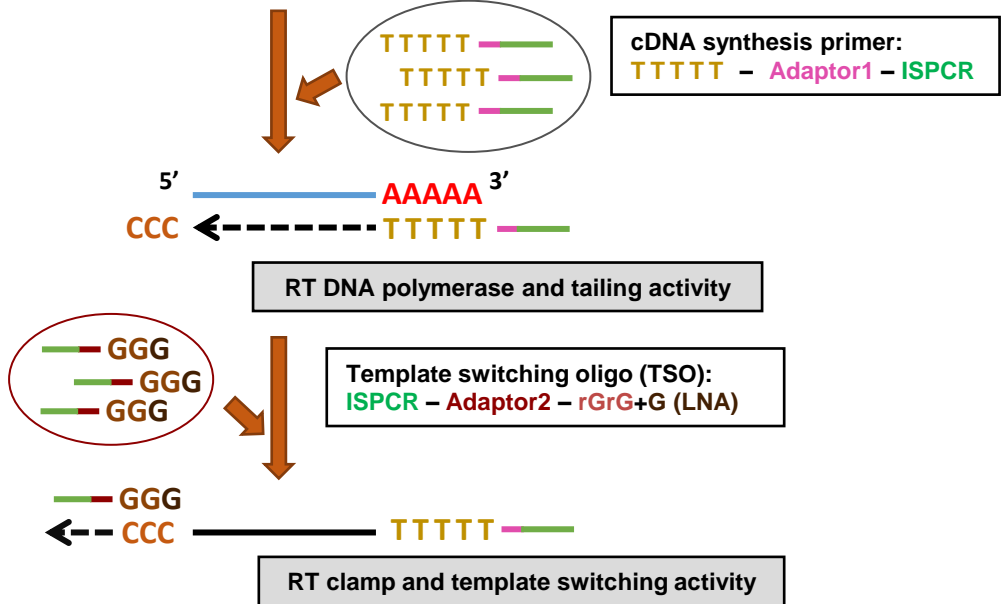
2. RNA isolation



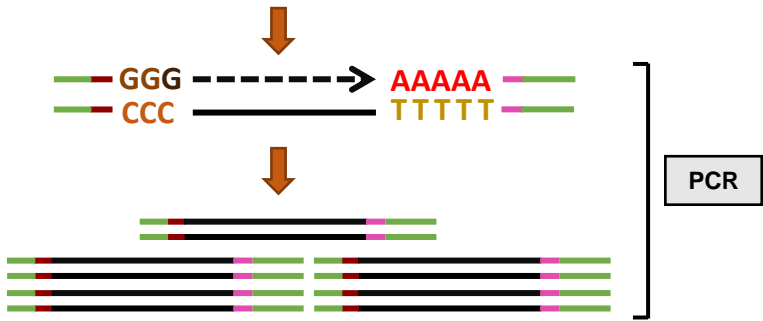
3. RNA depletion or selection



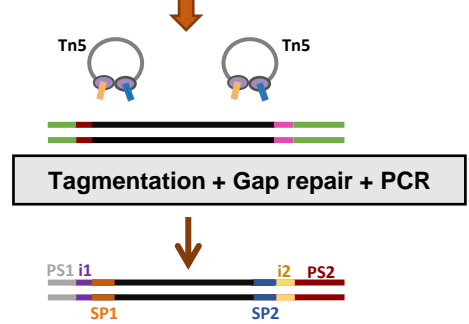
4. Reverse transcription: first strand cDNA synthesis



5. Second strand DNA synthesis and amplification



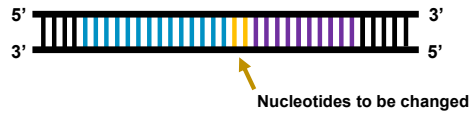
6. Library construction



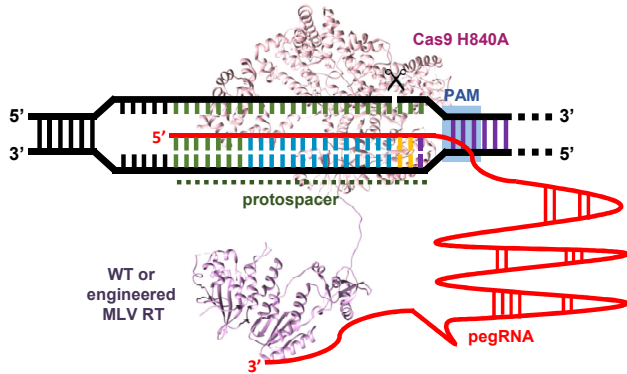
7. Sequencing and analysis



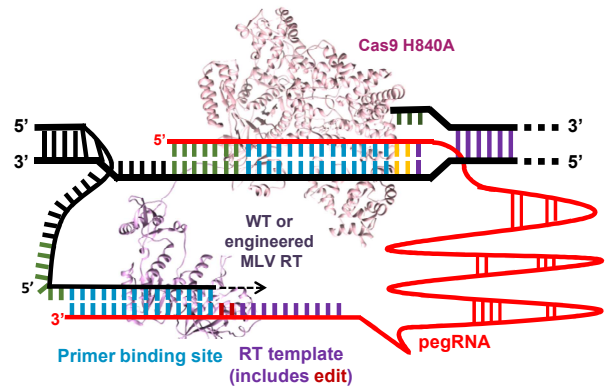
1. Target DNA



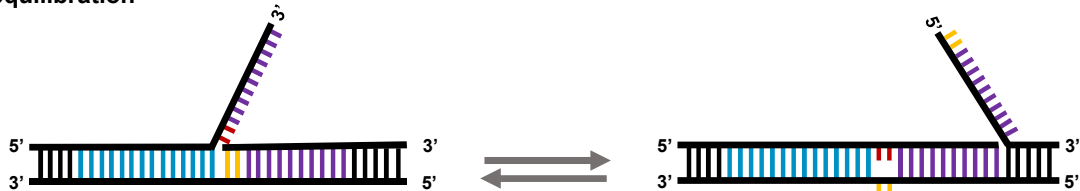
2. Hybridization to target DNA and single-strand cleavage by Cas9 H840A



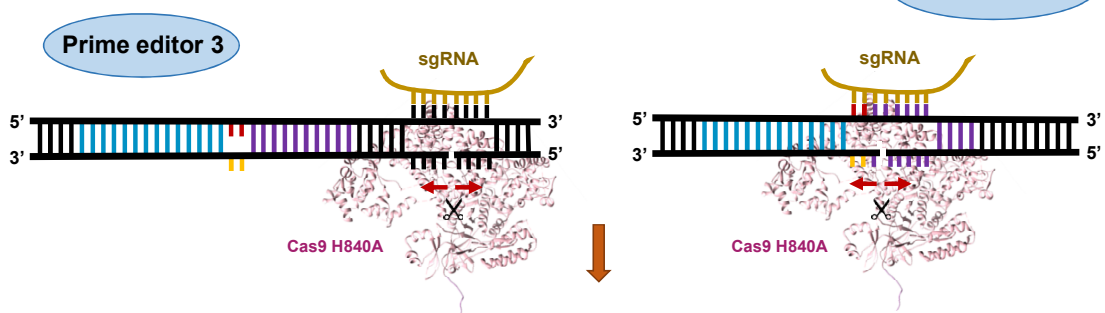
3. Hybridization to the primer binding site and reverse transcription



4. Flap equilibration



5. Nicking and correcting the unedited strand



6. Mismatch repair and final edited DNA



Supporting information

Reverse transcriptase: from transcriptomics to genome editing

Samara Martín-Alonso, Estrella Frutos-Beltrán and Luis Menéndez-Arias *

Centro de Biología Molecular “Severo Ochoa” (Consejo Superior de Investigaciones Científicas and Universidad Autónoma de Madrid), c/ Nicolás Cabrera 1, Campus de Cantoblanco-UAM, 28049 Madrid, Spain

* Corresponding author E-mail: lmendez@cbm.csic.es

S.M.-A. and E.F.-B. should be considered joint first authors

Table S1. Characteristics and applications of commercially available RTs (References)

Enzymes	References
<p>AMV Reverse Transcriptase (New England Biolabs, Promega, Clontech-Takara) (WT AMV RT)</p>	<ol style="list-style-type: none"> (1) New England Biolabs (https://international.neb.com/products/m0277-amv-reverse-transcriptase#Product%20Information) (2) Promega (https://www.promega.es/products/pcr/rt-pcr/amv-reverse-transcriptase/?catNum=M5101) (3) Clontech-Takara (https://www.takarabio.com/products/cdna-synthesis/reverse-transcriptases/amv-reverse-transcriptase-xl) (4) Deryusheva, S. and Gall, J.G. (2017). Dual nature of pseudouridylation in U2 snRNA: Pus1p-dependent and Pus1p-independent activities in yeasts and higher eukaryotes. <i>RNA</i> 23, 1060-1067 (5) Hizi, A. and Herschhorn, A. (2008). Retroviral reverse transcriptases (other than those of HIV-1 and murine leukemia virus): a comparison of their molecular and biochemical properties. <i>Virus Res.</i> 134, 203-220 (6) Werner, S. <i>et al.</i> (2020). Machine learning of reverse transcription signatures of variegated polymerases allows mapping and discrimination of methylated purines in limited transcriptomes. <i>Nucleic Acids Res.</i> 48, 3734-3746
<p>eAMV™ (Sigma-Merck) (AMV RT)</p>	<ol style="list-style-type: none"> (1) Sigma-Merck (https://www.sigmaaldrich.com/catalog/product/sigma/a4464?lang=es&region=ES) (2) Zucha, D. <i>et al.</i> (2020). Performance comparison of reverse transcriptases for single-cell studies. <i>Clin. Chem.</i> 66, 217-228
<p>ThermoScript™ (Invitrogen–Thermo Fisher Scientific) * (AMV RT)</p>	<ol style="list-style-type: none"> (1) Invitrogen-Thermo Fisher Scientific (https://www.thermofisher.com/order/catalog/product/11731023?SID=srch-srp-11731023#/11731023?SID=srch-srp-11731023) (2) Lelyveld, V.S. <i>et al.</i> (2019). DNA polymerase activity on synthetic N3'→ P5' phosphoramidate DNA templates. <i>Nucleic Acids Res.</i> 47, 8941-8949 (3) Markou, A. <i>et al.</i> (2008). Prognostic value of mature microRNA-21 and microRNA-205 overexpression in non–small cell lung cancer by quantitative real-time RT-PCR. <i>Clin. Chem.</i> 54, 1696-1704
<p>M-MLV Reverse transcriptase (Invitrogen-Thermo Fisher Scientific, New England Biolabs, Promega, Sigma-Merck) (WT MLV RT)</p>	<ol style="list-style-type: none"> (1) Invitrogen-Thermo Fisher Scientific (https://www.thermofisher.com/order/catalog/product/28025013#/28025013) (2) New England Biolabs (https://international.neb.com/products/m0253-m-mulv-reverse-transcriptase#Product%20Information) (3) Promega (https://www.promega.es/products/pcr/rt-pcr/m-mlv-reverse-transcriptase/?catNum=M1701) (4) Sigma-Merck (https://www.sigmaaldrich.com/catalog/product/sigma/m1302?lang=es&region=ES&gclid=EAlaIqo bChMlpKiCjamu6QIVFJzVCh0C5Ag-EAAYASAAEgLSwFD BwE) (5) Bagnoli, J.W. <i>et al.</i> (2018). Sensitive and powerful single-cell RNA sequencing using mcSCRBS-seq. <i>Nat. Commun.</i> 9, 1-8 (6) Coté, M.L. and Roth, M.J. (2008). Murine leukemia virus reverse transcriptase: structural comparison with HIV-1 reverse transcriptase. <i>Virus Res.</i> 134, 186-202. (7) Werner, S. <i>et al.</i> (2020). Machine learning of reverse transcription signatures of variegated polymerases allows mapping and discrimination of methylated purines in limited transcriptomes. <i>Nucleic Acids Res.</i> 48, 3734-3746 (8) Zucha, D. <i>et al.</i> (2020). Performance comparison of reverse transcriptases for single-cell studies. <i>Clin. Chem.</i> 66, 217-228

Table S1 (continuation)

Enzymes	References
<p>GoScript™ (Promega) (WT MLV RT)</p>	<p>(1) Promega (https://www.promega.es/products/pcr/rt-pcr/goscript-reverse-transcriptase/?catNum=A5003)</p> <p>(2) Bagnoli, J.W. <i>et al.</i> (2018). Sensitive and powerful single-cell RNA sequencing using mcSCRB-seq. <i>Nat. Commun.</i> 9, 1-8</p> <p>(3) Jang, M.K. <i>et al.</i> (2016). RNA-Seq analysis reveals a negative role of KLF16 in adipogenesis. <i>PLoS One</i>, 11, e0162238</p> <p>(4) Lin, X. <i>et al.</i> (2014). Increased expression of long noncoding RNA ABHD11-AS1 in gastric cancer and its clinical significance. <i>Med. Oncol.</i> 31, 42.</p> <p>(5) Shi, D.L. <i>et al.</i> (2016). MicroRNA-27a inhibits cell migration and invasion of fibroblast-like synoviocytes by targeting follistatin-like protein 1 in rheumatoid arthritis. <i>Mol. Cells</i> 39, 611.</p> <p>(6) Werner, S. <i>et al.</i> (2020). Machine learning of reverse transcription signatures of variegated polymerases allows mapping and discrimination of methylated purines in limited transcriptomes. <i>Nucleic Acids Res.</i> 48, 3734-3746</p>
<p>MultiScribe™ (Thermo Fisher Scientific) (WT MLV RT)</p>	<p>(1) Thermo Fisher Scientific. (https://www.thermofisher.com/order/catalog/product/4311235#/4311235)</p> <p>(1) Chettri, J.K. <i>et al.</i> (2011). PAMP induced expression of immune relevant genes in head kidney leukocytes of rainbow trout (<i>Oncorhynchus mykiss</i>). <i>Dev. Comp. Immunol.</i> 35, 476-482</p>
<p>AccuScript Hi-Fi (Stratagene-Agilent) (MLV RT)</p>	<p>(2) Stratagene-Agilent (https://www.agilent.com/en/product/real-time-pcr-(qpcr)/real-time-pcr-(qpcr)-cdna-synthesis-kits/accuscript/accuscript-hi-fi-reverse-transcriptase-232708)</p> <p>(3) Werner, S. <i>et al.</i> (2020). Machine learning of reverse transcription signatures of variegated polymerases allows mapping and discrimination of methylated purines in limited transcriptomes. <i>Nucleic Acids Res.</i> 48, 3734-3746</p> <p>(4) Zucha, D. <i>et al.</i> (2020). Performance comparison of reverse transcriptases for single-cell studies. <i>Clin. Chem.</i> 66, 217-228</p>
<p>Affinity Script (Agilent) (MLV RT)</p>	<p>(1) Agilent (https://www.agilent.com/en/product/real-time-pcr-(qpcr)/real-time-pcr-(qpcr)-cdna-synthesis-kits/affinityscript/affinityscript-multiple-temperature-reverse-transcriptase-232706#productdetails)</p> <p>(2) Hogrefe, H. <i>et al.</i> (2008) Mutant reverse transcriptase and methods of use. European Patent EP1931772B1. WO 2007/022045 (22.02.2007 Gazette 2007/08) (https://patents.google.com/patent/EP1931772B1/en)</p> <p>(3) Arezi, B. and Hogrefe, H. (2009). Novel mutations in Moloney murine leukemia virus reverse transcriptase increase thermostability through tighter binding to template-primer. <i>Nucleic Acids Res.</i> 37, 473-481</p> <p>(4) Barker, M.S. <i>et al.</i> (2009). Paleopolyploidy in the Brassicales: analyses of the Cleome transcriptome elucidate the history of genome duplications in Arabidopsis and other Brassicales. <i>Genome Biol. Evol.</i> 1, 391-399</p> <p>(5) Rotem, A. <i>et al.</i> (2015). High-throughput single-cell labeling (Hi-SCL) for RNA-seq using drop-based microfluidics. <i>PLoS One</i> 10, e0116328</p> <p>(6) Werner, S. <i>et al.</i> (2020). Machine learning of reverse transcription signatures of variegated polymerases allows mapping and discrimination of methylated purines in limited transcriptomes. <i>Nucleic Acids Res.</i> 48, 3734-3746</p>
<p>ArrayScript™ (Invitrogen-Ambion-Thermo Fisher Scientific) (MLV RT)</p>	<p>(1) Invitrogen-Ambion-Thermo Fisher Scientific (https://www.fishersci.es/shop/products/ambion-arrayscript-reverse-transcriptase/10301275)</p> <p>(2) Hashimshony, T. (2016) CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. <i>Genome Biol.</i> 17, 77</p>

Table S1 (continuation)

Enzymes	References
BioScript™ (Bioline) (MLV RT)	(1) Bioline (https://www.bioline.com/de/bioscript.html) (2) Barker, M.S. <i>et al.</i> (2009). Paleopolyploidy in the Brassicales: analyses of the Cleome transcriptome elucidate the history of genome duplications in Arabidopsis and other Brassicales. <i>Genome Biol. Evol.</i> 1, 391-399
CycleScript™ (Bioneer) (MLV RT)	(1) Bioneer (https://us.bioneer.com/products/enzyme/enzymeCycleScriptReversetechnical.aspx)
EnzScript™ (Biozym) (MLV RT)	(1) Biozym (https://www.biozym.com/DesktopModules/Webshop/shopdisplayproducts.aspx?productid=29617&id=2355) (2) Bagnoli, J.W. <i>et al.</i> (2018). Sensitive and powerful single-cell RNA sequencing using mcSCRBS-seq. <i>Nat. Commun.</i> 9, 1-8
EpiScript™ (Lucigen-Biosearch Technologies) (MLV RT)	(1) (Lucigen) Biosearch Technologies (https://www.lucigen.com/EpiScript-RNase-H-Reverse-Transcriptase/) (2) Gene Target Solutions (https://www.genetargetsolutions.com.au/product/reverse-transcriptase-selection-guide-3/) (3) Deryusheva, S. and Gall, J.G. (2017). Dual nature of pseudouridylation in U2 snRNA: Pus1p-dependent and Pus1p-independent activities in yeasts and higher eukaryotes. <i>RNA</i> 23, 1060-1067 (4) Werner, S. <i>et al.</i> (2020). Machine learning of reverse transcription signatures of variegated polymerases allows mapping and discrimination of methylated purines in limited transcriptomes. <i>Nucleic Acids Res.</i> 48, 3734-3746
Expand™ Reverse Transcriptase (Roche-Merck) (MLV RT)	(1) Roche-Merck (https://www.sigmaaldrich.com/catalog/product/roche/ertro?lang=es&region=ES)
FIREScript (Solis Biodyne) (MLV RT)	(1) Solis Biodyne (https://www.solisbiodyne.com/EN/products/cdna_synthesis_reverse_transcription/firescript_rt/)
GrandScript (TATAA Biocenter) * (MLV RT)	(1) TATAA Biocenter (https://webshop.tataa.com/product.html/tataa-grandscript-cdna-synthesis-kit) (2) Åman, P. <i>et al.</i> (2016). Regulatory mechanisms, expression levels and proliferation effects of the FUS-DDIT3 fusion oncogene in liposarcoma. <i>J. Pathol.</i> 238, 689-699
iScript™ (BioRad) * (MLV RT)	(1) BioRad (https://www.bio-rad.com/es-es/SearchResults?Text=iscrpt) (2) Zucha, D. <i>et al.</i> (2020). Performance comparison of reverse transcriptases for single-cell studies. <i>Clin. Chem.</i> 66, 217-228
Maxima™ RT (Thermo Fisher Scientific) (MLV RT)	(1) Thermo Fisher Scientific. (https://www.thermofisher.com/es/es/home/brands/thermo-scientific/molecular-biology/thermo-scientific-reverse-transcription/maxima.html) (2) Bagnoli, J.W. <i>et al.</i> (2018). Sensitive and powerful single-cell RNA sequencing using mcSCRBS-seq. <i>Nat. Commun.</i> 9, 1-8 (3) Dickel, D.E. <i>et al.</i> (2018). Ultraconserved enhancers are required for normal development. <i>Cell</i> 172, 491-499.e15 (4) Semrau, S. <i>et al.</i> (2017). Dynamics of lineage commitment revealed by single-cell transcriptomics of differentiating embryonic stem cells. <i>Nature Commun.</i> 8, 1096 (5) Zucha, D. <i>et al.</i> (2020). Performance comparison of reverse transcriptases for single-cell studies. <i>Clin. Chem.</i> 66, 217-228

Table S1 (continuation)

Enzymes	References
<p>MonsterScript™ (Epicentre-Ecogen) (MLV RT)</p>	<p>(1) Epicentre-Ecogen (http://www.ecogen.com/producto.asp?id=425) (2) Gene target solutions (https://www.genetargetsolutions.com.au/product/reverse-transcriptase-selection-guide-3/) (3) Tebu-Bio (https://www.tebu-bio.com/blog/2015/03/31/expired-patents-may-mean-lower-prices-for-superscript-ii-reverse-transcriptase-users/) (4) Tang, C. <i>et al.</i> (2018). Template switching causes artificial junction formation and false identification of circular RNAs. <i>bioRxiv</i>, 259556 (5) Werner, S. <i>et al.</i> (2020). Machine learning of reverse transcription signatures of variegated polymerases allows mapping and discrimination of methylated purines in limited transcriptomes. <i>Nucleic Acids Res.</i> 48, 3734-3746</p>
<p>PrimeScript™ (Takara) (MLV RT)</p>	<p>(1) Takara (https://www.takarabio.com/search-results?term=2680A) (2) Barker, M.S. <i>et al.</i> (2009). Paleopolyploidy in the Brassicales: analyses of the Cleome transcriptome elucidate the history of genome duplications in Arabidopsis and other Brassicales. <i>Genome Biol. Evol.</i> 1, 391-399 (3) Coenen-Stass, A.M.L. <i>et al.</i> (2018). Evaluation of methodologies for microRNA biomarker detection by next generation sequencing. <i>RNA Biol.</i> 15, 1133–1145 (4) Zucha, D. <i>et al.</i> (2020). Performance comparison of reverse transcriptases for single-cell studies. <i>Clin. Chem.</i> 66, 217-228</p>
<p>PrimeScript™ II (Takara) (MLV RT)</p>	<p>(1) Takara (https://www.takarabio.com/search-results?term=2690A) (2) Keren-Shaul, H. <i>et al.</i> (2019). MARS-seq2.0: an experimental and analytical pipeline for indexed sorting combined with single-cell RNA sequencing. <i>Nat. Protoc.</i> 14, 1841-1862</p>
<p>PrimeScript™ III (Takara) * (MLV RT)</p>	<p>(1) Takara (https://www.takarabio.com/products/real-time-pcr/real-time-pcr-kits/qpcr-with-probe-detection/one-step-primerscript-iii-rt-pcr-kit)</p>
<p>PrimeScript™ IV (Takara) * (MLV RT)</p>	<p>(1) Takara (https://www.takarabio.com/products/cdna-synthesis/cdna-synthesis-kits/primerscript-cdna-synthesis-kits/primerscript-iv-first-strand-cdna-synthesis-mix)</p>
<p>ProtoScript® (New England Biolabs) * (MLV RT)</p>	<p>(1) New England Biolabs (https://international.neb.com/products/e6300-protoscript-first-strand-cdna-synthesis-kit#Product%20Information) (2) Kureel, A.K. <i>et al.</i> (2019). Identification of a novel transcript variant of the human CD6 gene that lacks exon 9. <i>Immunobiol.</i> 224, 666-671</p>
<p>ProtoScript® II (New England Biolabs) (MLV RT)</p>	<p>(1) New England Biolabs (https://international.neb.com/products/m0368-protoscript-ii-reverse-transcriptase#Product%20Information) (2) Bagnoli, J.W. <i>et al.</i> (2018). Sensitive and powerful single-cell RNA sequencing using mcSCRB-seq. <i>Nat. Commun.</i> 9, 1-8 (5) Coenen-Stass, A.M.L. <i>et al.</i> (2018). Evaluation of methodologies for microRNA biomarker detection by next generation sequencing. <i>RNA Biol.</i> 15, 1133–1145 (3) Lelyveld, V.S. <i>et al.</i> (2019). DNA polymerase activity on synthetic N3'→ P5' phosphoramidate DNA templates. <i>Nucleic Acids Res.</i> 47, 8941-8949 (4) Liu, Y. <i>et al.</i> (2020). Direct detection of circRNA in real samples using reverse transcription-rolling circle amplification. <i>Anal. Chim. Acta</i> 1101, 169-175 (5) Potapov, V. <i>et al.</i> (2018). Base modifications affecting RNA polymerase and reverse transcriptase fidelity. <i>Nucleic Acids Res.</i> 46, 5753-5763 (6) Werner, S. <i>et al.</i> (2020). Machine learning of reverse transcription signatures of variegated polymerases allows mapping and discrimination of methylated purines in limited transcriptomes. <i>Nucleic Acids Res.</i> 48, 3734-3746</p>

Table S1 (continuation)

Enzymes	References
<p>qScript (Quantabio) * (MLV RT)</p>	<p>(1) Quatabio (https://www.quantabio.com/resources?q=qscript;https://www.quantabio.com/media/contenttype/IFU-105.1_REV_02_95161_qScript_XLT_cDNA_SuperMix.pdf) (2) Zucha, D. <i>et al.</i> (2020). Performance comparison of reverse transcriptases for single-cell studies. <i>Clin. Chem.</i> 66, 217-228</p>
<p>RevertAid™ (Thermo Fisher Scientific) (MLV RT)</p>	<p>(1) Thermo Fischer Scientific (https://www.thermofisher.com/order/catalog/product/EP0441#/EP0441) (2) Janulaitis, A. <i>et al.</i> (2014) Production of nucleic acid. U.S. patent no. US 8,835,148 B2 (https://patents.google.com/patent/US8835148B2/en) (3) Bagnoli, J.W. <i>et al.</i> (2018). Sensitive and powerful single-cell RNA sequencing using mcSCRBS-seq. <i>Nat. Commun.</i> 9, 1-8 (4) Chen, X. <i>et al.</i> (2018). Efficient in situ barcode sequencing using padlock probe-based BaristaSeq. <i>Nucleic Acids Res.</i> 46, e22 (5) Werner, S. <i>et al.</i> (2020). Machine learning of reverse transcription signatures of variegated polymerases allows mapping and discrimination of methylated purines in limited transcriptomes. <i>Nucleic Acids Res.</i> 48, 3734-3746</p>
<p>ReverTra Ace® (Toyobo) (MLV RT)</p>	<p>(1) Toyobo (https://www.toyobo-global.com/seihin/xr/lifescience/products/mod_001.html) (2) Wang, Z. <i>et al.</i> (2019). Comparative transcriptome profiling of resistant and susceptible sugarcane genotypes in response to the airborne pathogen <i>Fusarium verticillioides</i>. <i>Mol. Biol. Rep.</i> 46, 3777-3789</p>
<p>RevertUp II™ (Biotechrabbit) (MLV RT)</p>	<p>(1) Biotechrabbit (https://www.biotechrabbit.com/products/endpoint-pcr/reverse-transcription-and-rt-pcr/revertuptm-ii-reverse-transcriptase-200-u-l.html) (2) Bagnoli, J.W. <i>et al.</i> (2018). Sensitive and powerful single-cell RNA sequencing using mcSCRBS-seq. <i>Nat. Commun.</i> 9, 1-8</p>
<p>Rocketscript™ (Bioneer) (MLV RT)</p>	<p>(1) Bioneer (https://us.bioneer.com/products/enzyme/rocketscript-rt-overview.aspx;https://us.bioneer.com/products/enzyme/rocketscript-rt-RNase-overview.aspx) (2) Sun, Y. <i>et al.</i> (2019). Overexpression of TaBADH increases salt tolerance in Arabidopsis. <i>Can. J. Plant Sci.</i> 99, 546-555</p>
<p>Script (Jena Bioscience) (MLV RT)</p>	<p>(1) Jena Bioscience (https://www.jenabioscience.com/molecular-biology/reverse-transcription/reverse-transcriptases/pcr-505-script-reverse-transcriptase;https://www.jenabioscience.com/molecular-biology/reverse-transcription/one-step-rt-pcr/pcr-509-script-one-step-rt-pcr-kit)</p>
<p>SMART® (Clontech-Takara) (MLV RT)</p>	<p>(1) Clontech-Takara (https://www.takarabio.com/products/cdna-synthesis/reverse-transcriptases/smart-mmlv-rt?catalog=639522)</p>
<p>SMARTScribe™ (Clontech-Takara) (MLV RT)</p>	<p>(1) Clontech-Takara (https://www.takarabio.com/products/cdna-synthesis/reverse-transcriptases/smartscribe) (2) Akan, I. <i>et al.</i> (2016) Drosophila O-GlcNAcase deletion globally perturbs chromatin O-GlcNAcylation. <i>J. Biol. Chem.</i> 291, 9906-9919 (3) Bagnoli, J.W. <i>et al.</i> (2018). Sensitive and powerful single-cell RNA sequencing using mcSCRBS-seq. <i>Nat. Commun.</i> 9, 1-8 (4) Zucha, D. <i>et al.</i> (2020). Performance comparison of reverse transcriptases for single-cell studies. <i>Clin. Chem.</i> 66, 217-228</p>

Table S1 (continuation)

Enzymes	References
<p>SuperScript™ II (Invitrogen-Thermo Fisher Scientific) (MLV RT)</p>	<ol style="list-style-type: none"> (1) Invitrogen-Thermo Fisher Scientific (https://www.thermofisher.com/order/catalog/product/18064014?SID=srch-srp-18064014#/18064014?SID=srch-srp-18064014) (2) Kotewicz, M.L. and Gerard, G.F. (1993). Cloned genes encoding reverse transcriptase lacking RNase H activity. US patent no. 5,244,797 B1 (https://patents.google.com/patent/US5244797B1/en) (3) Bagnoli, J.W. <i>et al.</i> (2018). Sensitive and powerful single-cell RNA sequencing using mcSCRBS-seq. <i>Nat. Commun.</i> 9, 1-8 (4) Barker, M.S. <i>et al.</i> (2009). Paleopolyploidy in the Brassicales: analyses of the <i>Cleome</i> transcriptome elucidate the history of genome duplications in <i>Arabidopsis</i> and other Brassicales. <i>Genome Biol. Evol.</i> 1, 391-399 (5) Chaput, J.C. <i>et al.</i> (2003). DNA polymerase-mediated DNA synthesis on a TNA template. <i>J. Amer. Chem. Soc.</i> 125, 856-857 (6) Chen, Y.S. <i>et al.</i> (2019). 5-Methylcytosine analysis by RNA-BisSeq. <i>Methods Mol. Biol.</i> 1870, 237-248 (7) Levesque-Sergerie, J.P. <i>et al.</i> (2007). Detection limits of several commercial reverse transcriptase enzymes: impact on the low- and high-abundance transcript levels assessed by quantitative RT-PCR. <i>BMC Mol. Biol.</i> 8, 93 (8) Zucha, D. <i>et al.</i> (2020). Performance comparison of reverse transcriptases for single-cell studies. <i>Clin. Chem.</i> 66, 217-228
<p>SuperScript™ III (Invitrogen-Thermo Fisher Scientific) (MLV RT)</p>	<ol style="list-style-type: none"> (1) Invitrogen-Thermo Fisher Scientific (https://www.thermofisher.com/order/catalog/product/18080093?SID=srch-srp-18080093#/18080093?SID=srch-srp-18080093) (2) Potter, R.J. and Rosenthal, K. (2006). High fidelity reverse transcriptases and uses thereof. US patent no. 7,056,716 B2 (https://patents.google.com/patent/US7056716B2/en) (3) Crouzier, L. <i>et al.</i> (2012). Efficient reverse transcription using locked nucleic acid nucleotides towards the evolution of nuclease resistant RNA aptamers. <i>PLoS One</i>, 7, e35990 (4) Sheng, K. <i>et al.</i> (2017). Effective detection of variation in single-cell transcriptomes using MATQ-seq. <i>Nat. Methods</i>, 14, 267 (5) Werner, S. <i>et al.</i> (2020). Machine learning of reverse transcription signatures of variegated polymerases allows mapping and discrimination of methylated purines in limited transcriptomes. <i>Nucleic Acids Res.</i> 48, 3734-3746 (6) Zucha, D. <i>et al.</i> (2020). Performance comparison of reverse transcriptases for single-cell studies. <i>Clin. Chem.</i> 66, 217-228
<p>SuperScript™ IV (Invitrogen-Thermo Fisher Scientific) (MLV RT)</p>	<ol style="list-style-type: none"> (1) Invitrogen-Thermo Fisher Scientific (https://www.thermofisher.com/order/catalog/product/18090050#/18090050) (2) Rogers, J. and Potter, R. (2018). Novel reverse transcriptases for use in high temperature nucleic acid synthesis. US patent application US 2018/0010105 A1 (https://patents.google.com/patent/US20180010105) (3) Kietrys, A.M. <i>et al.</i> (2017). Fingerprints of modified RNA bases from deep sequencing profiles. <i>J. Amer. Chem. Soc.</i> 139, 17074-17081 (4) Werner, S. <i>et al.</i> (2020). Machine learning of reverse transcription signatures of variegated polymerases allows mapping and discrimination of methylated purines in limited transcriptomes. <i>Nucleic Acids Res.</i> 48, 3734-3746 (5) Zucha, D. <i>et al.</i> (2020). Performance comparison of reverse transcriptases for single-cell studies. <i>Clin. Chem.</i> 66, 217-228

Table S1 (continuation)

Enzymes	References
<p>Transcriptor Reverse Transcriptase (Roche-Merck) (MLV RT)</p>	<p>(1) Roche-Merck. (https://www.sigmaaldrich.com/catalog/product/roche/transrtro?lang=es&region=ES)</p> <p>(2) Kubista, M. et al. (2014) Lysis and reverse transcription for mRNA quantification. US patent no. 8,623,602 B2 (https://patents.google.com/patent/US8623602B2/en)</p> <p>(3) Brooke-Powell, E.T. et al. (2004). Use of transcriptor reverse transcriptase in microarray analysis. <i>Biochemica-Mannheim- 1</i>, 27-30</p> <p>(4) Stanitz, E. et al. (2013). Evaluation of MicroRNA expression pattern of gastric adenocarcinoma associated with socioeconomic, environmental and lifestyle factors in northwestern Hungary. <i>Anticancer Res.</i> 33, 3195-3200</p>
<p>HIV-1 RT (Abcam, Ambion-Thermo Fisher Scientific, Worthington) (HIV-1 RT)</p>	<p>(1) Abcam (https://www.abcam.com/recombinant-hiv1-reverse-transcriptase-protein-active-ab63979.html)</p> <p>(2) Ambion-Invitrogen-Thermo Fischer Scientific (https://www.thermofisher.com/es/es/home/references/ambion-tech-support/probe-labeling-systems/tech-notes/recombinant-hiv-reverse-transcriptase.html)</p> <p>(3) Worthington (http://www.worthington-biochem.com/HIVRT/cat.html)</p> <p>(4) Kempeneers, V. et al. (2005). Investigation of the DNA-dependent cyclohexenyl nucleic acid polymerization and the cyclohexenyl nucleic acid-dependent DNA polymerization. <i>Nucleic Acids Res.</i> 33, 3828-3836</p> <p>(5) Sinha, S. et al. (2004). 2', 5'-Linked DNA Is a Template for Polymerase-Directed DNA Synthesis. <i>J. Amer. Chem. Soc.</i> 126, 40-41</p> <p>(6) Werner, S. et al. (2020). Machine learning of reverse transcription signatures of variegated polymerases allows mapping and discrimination of methylated purines in limited transcriptomes. <i>Nucleic Acids Res.</i> 48, 3734-3746</p>
<p>Biotoools High Retrotranscriptase (Biotoools) (HIV-1 RT)</p>	<p>(1) Biotoools (http://biotoools.eu/es/cdna/347-10077-4116-Biotoools-High-Retrotranscriptase.html)</p> <p>(2) Barrioluengo, V. et al. (2011) Thermostable HIV-1 group O reverse transcriptase variants with the same fidelity as murine leukaemia virus reverse transcriptase. <i>Biochem. J.</i> 436, 599-607</p>
<p>Sunscript® (4Basebio) (HIV-1 RT)</p>	<p>(1) 4Basebio. Genomics product guide, pp. 22-26 (https://www.4basebio.com/)</p> <p>(2) Menéndez-Arias, L. et al. (2016). HIV type 1 group O reverse transcriptases that are active at high temperatures. US patent no. 9.428,738 B2 (https://patents.google.com/patent/US9428738B2)</p> <p>(3) Luczkowiak, J. et al. (2018). Template-primer binding affinity and RNase H cleavage specificity contribute to the strand transfer efficiency of HIV-1 reverse transcriptase. <i>J. Biol. Chem.</i> 293, 13351-13363</p> <p>(4) Matamoros, T. et al. (2013) Major groove binding track residues of the connection subdomain of human immunodeficiency virus type 1 reverse transcriptase enhance cDNA synthesis at high temperatures. <i>Biochemistry</i> 52, 9318-9328</p>
<p>Marathon RT (Kerafast) (Intron RT)</p>	<p>(1) Kerafast (https://www.kerafast.com/productgroup/855/marathonrt-reverse-transcriptase)</p> <p>(2) Guo, L.T. et al. (2020). Sequencing and structure probing of long RNAs using MarathonRT: A next-generation reverse transcriptase. <i>J. Mol. Biol.</i> 432, 3338-3352</p> <p>(3) Zhao, C. et al. (2018). An ultraproccessive, accurate reverse transcriptase encoded by a metazoan group II intron. <i>RNA</i> 24, 183-195</p>

Table S1 (continuation)

Enzymes	References
<p>TGIRT®-III RT (InGex) (Intron RT)</p>	<p>(1) InGex (https://www.ingex.com/tgirt-enzyme/; http://www.ingex.com/content/Total%2ORNA_Kit_012317.pdf) (2) Mohr, S. <i>et al.</i> (2013). Thermostable group II intron reverse transcriptase fusion proteins and their use in cDNA synthesis and next-generation RNA sequencing. <i>RNA</i> 19, 958-970 (3) Qin, Y. <i>et al.</i> (2016). High-throughput sequencing of human plasma RNA by using thermostable group II intron reverse transcriptases. <i>RNA</i> 22, 111-128 (4) Werner, S. <i>et al.</i> (2020). Machine learning of reverse transcription signatures of variegated polymerases allows mapping and discrimination of methylated purines in limited transcriptomes. <i>Nucleic Acids Res.</i> 48, 3734-3746 (5) Wu, D.C. and Lambowitz, A.M. (2017). Facile single-stranded DNA sequencing of human plasma DNA via thermostable group II intron reverse transcriptase template switching. <i>Sci. Rep.</i> 7, 8421. Erratum in: <i>Sci Rep.</i> (2018) 8, 4056 (6) Zhao, C. <i>et al.</i> (2018). An ultraprocessive, accurate reverse transcriptase encoded by a metazoan group II intron. <i>RNA</i> 24, 183-195</p>
<p>BcaBEST Polymerase (Takara) * (Bacterial DNA pol)</p>	<p>(1) Takara (https://www.takarabio.com/products/cdna-synthesis/cdna-synthesis-kits/high-gc-or-secondary-structure?catalog=RR023A) (2) Enoki, T. <i>et al.</i> (2006). Isothermal chimeric primer nucleic acid amplification methods using blocking ooglionucleotide. US patent no. 7,056,671 B2 (https://patents.google.com/patent/US7056671B2/en)</p>
<p>Bst 3.0 DNA Polymerase (New England Biolabs) (Bacterial DNA pol)</p>	<p>(1) New England Biolabs (https://international.neb.com/products/m0374-bst-3-0-dna-polymerase#Product%20Information) (2) Ağel, H.E. <i>et al.</i> (2020). Optimization of the Isothermal Amplification Method for Mycobacterium tuberculosis Detection and Visualization for Fieldwork. <i>Turk. J. Med. Sci.</i>, doi: 10.3906/sag-1910-6 (3) Garafutdinov, R.R. <i>et al.</i> (2020). The influence of reaction conditions on DNA multimerization during isothermal amplification with Bst exo- DNA polymerase. <i>Appl. Biochem. Biotechnol.</i> 190, 758-771</p>
<p>RapiDxFire™ Reverse Transcriptase (Lucigen-Biosearch Technologies) (Bacterial DNA pol)</p>	<p>(1) Lucigen-Biosearch Technologies (https://www.lucigen.com/RapiDxFire-Thermostable-Reverse-Transcriptase/)</p>
<p>Volcano2G DNA polymerase (myPOLS Biotec) (Bacterial DNA pol)</p>	<p>(1) myPOLS Biotec (https://www.mypols.de/products/volcano2g-dna-polymerase) (2) Blatter, N. <i>et al.</i> (2013). Structure and function of an RNA-reading thermostable DNA polymerase. <i>Angew. Chem. Int. Ed. Engl.</i> 52, 11935-11939 (3) Chovancova, P. <i>et al.</i> (2017). Reverse-transcription quantitative PCR directly from cells without RNA extraction and without isothermal reverse-transcription: a 'zero-step' RT-qPCR protocol. <i>Biol. Method Protoc.</i> 2, bpx008 (4) Werner, S. <i>et al.</i> (2020). Machine learning of reverse transcription signatures of variegated polymerases allows mapping and discrimination of methylated purines in limited transcriptomes. <i>Nucleic Acids Res.</i> 48, 3734-3746</p>
<p>Volcano3G DNA polymerase (myPOLS Biotec) * (Bacterial DNA pol)</p>	<p>(1) myPOLS Biotec (https://www.mypols.de/products/volcano3g-rt-pcr-probe-2x-master-mix)</p>
<p>SOLIScript (Solis Biodyne)</p>	<p>(1) Solis Biodyne (https://www.solisbiodyne.com/EN/products/cdna_synthesis_reverse_transcription/soliscrpt_rt)</p>

Table S1 (continuation)

Enzymes	References
Omniscript® (Qiagen) *	<p>(1) Qiagen (https://www.qiagen.com/us/products/discovery-and-translational-research/pcr-qpcr/pcr-enzymes-and-kits/reverse-transcription-and-cdna-synthesis-qpcr/omniscript-rt-kit/?clear=true#orderinginformation)</p> <p>(2) Vieira, A. <i>et al.</i> (2016). Comparative Validation of Conventional and RNA-Seq Data-Derived Reference Genes for qPCR Expression Studies of <i>Colletotrichum kahawae</i>. <i>PLoS One</i> 11, e0150651</p> <p>(3) Karsai, A. <i>et al.</i> (2002). Evaluation of a homemade SYBR green I reaction mixture for real-time PCR quantification of gene expression. <i>Biotechniques</i> 32, 790-792, 794-796</p>
SensiScript® (Qiagen) *	<p>(1) Qiagen (https://www.qiagen.com/us/products/discovery-and-translational-research/pcr-qpcr/pcr-enzymes-and-kits/reverse-transcription-and-cdna-synthesis-qpcr/sensiscript-rt-kit/?clear=true#orderinginformation)</p> <p>(2) Levesque-Sergerie, J.P. <i>et al.</i> (2007). Detection limits of several commercial reverse transcriptase enzymes: impact on the low- and high-abundance transcript levels assessed by quantitative RT-PCR. <i>BMC Mol. Biol.</i> 8, 93</p> <p>(3) Zucha, D. <i>et al.</i> (2020). Performance comparison of reverse transcriptases for single-cell studies. <i>Clin. Chem.</i> 66, 217-228</p>

Additional RTs included in kits for specific applications

Enzyme	References
3B High Retrotranscriptase (Indiamart)	(1) Indimart (https://www.indiamart.com/3b-blackbio/reverse-transcriptase-kits.html#3b-high-retrotranscriptase)
HotStaRT-Script Reverse (Qiagen)	(1) Qiagen (https://www.qiagen.com/us/products/discovery-and-translational-research/pcr-qpcr/real-time-pcr-enzymes-and-kits/probe-based-qpcr/quantinova-pathogen-ic-kit/?clear=true#orderinginformation)
ImProm™-II (Promega)	<p>(1) Promega. (https://www.promega.es/products/pcr/rt-pcr/improm-ii-reverse-transcriptase/?catNum=A3801)</p> <p>(2) Lv, G. <i>et al.</i> (2014). MicroRNA-451 regulates activating transcription factor 2 expression and inhibits liver cancer cell migration. <i>Oncol. Rep.</i> 32, 1021-1028</p> <p>(3) Reitmann, A. <i>et al.</i> (2017). Putative pathogenicity genes of <i>Phytophthora cinnamomi</i> identified via RNA-Seq analysis of pre-infection structures. <i>Eur. J. Plant Pathol.</i> 147, 211-228</p>
ipsogen® RT (Qiagen)	(1) Qiagen (https://www.qiagen.com/us/products/diagnostics-and-clinical-research/oncology/ipsogen-leukemia/ipsogen-rt-kit/?clear=true#orderinginformation)
Isotherm2G DNA polymerase (myPOLS Biotec)	(2) myPOLS Biotec (https://www.mypols.de/collections/rna-rt-pc/products/isotherm2g-dna-polymerase)
LunaScript™ (New England Biolabs)	(1) New England Biolabs (https://international.neb.com/products/e3010-luna-script-rt-supermix-kit#Product%20Information)(https://international.neb.com/products/pcr-qpcr-and-amplification-technologies/luna-qpcr-and-rt-qpcr-products/detail)
miScript RT (Qiagen)	(1) Qiagen (https://www.qiagen.com/de/products/discovery-and-translational-research/pcr-qpcr/qpcr-assays-and-instruments/mirna-qpcr-assay-and-panels/miscript-ii-rt-kit/?clear=true#orderinginformation)
NEXTFLEX® RNA-Seq Rapid Reverse transcriptase (Perkin Elmer)	(1) Perkinelmer (https://perkinelmer-appliedgenomics.com/home/products/library-preparation-kits/illumina-rna-library-prep-kits/nextflex-rapid-rna-seq-kit/)

Additional RTs included in kits for specific applications (continuation)

Enzyme	References
NEXTFLEX® RNA-Seq Rapid Reverse transcriptase 2.0 (Perkin Elmer)	(1) PerkinElmer (https://perkinelmer-appliedgenomics.com/home/products/library-preparation-kits/illumina-rna-library-prep-kits/nextflex-rapid-directional-rna-seq-kit-2-0/)
QuantiNova Reverse transcriptase (Qiagen)	(1) Qiagen (https://www.qiagen.com/dk/products/discovery-and-translational-research/pcr-qpcr/real-time-pcr-enzymes-and-kits/reverse-transcription-cdna-synthesis-qpcr/quantinova-reverse-transcription-kit/#orderinginformation)
Quantiscript Reverse transcriptase (Qiagen)	(1) Qiagen (https://www.qiagen.com/dk/products/discovery-and-translational-research/pcr-qpcr/real-time-pcr-enzymes-and-kits/reverse-transcription-cdna-synthesis-qpcr/quantitect-reverse-transcription-kit/?clear=true#orderinginformation)
RetroAMPTM RT DNA polymerase (Ecogen)	(1) Ecogen (http://www.ecogen.com/producto.asp?id=429) (2) Chang, L.S. <i>et al.</i> (2002). Differential expression of human 5S snoRNA genes. <i>Biochem. Biophys. Res. Commun.</i> 299, 196-200
SEQzyme (BioRad)	(1) BioRad (https://www.bio-rad.com/es-es/product/sequoia-complete-stranded-rna-library-prep-kit?ID=PZ1792E0801Y)
T-Script Enzyme (Qiagen)	(1) Qiagen (https://www.qiagen.com/it/products/discovery-and-translational-research/pcr-qpcr/real-time-pcr-enzymes-and-kits/reverse-transcription-cdna-synthesis-qpcr/quantitect-whole-transcriptome-kit/?clear=true#orderinginformation)
WarmStart RTx® Reverse transcriptase (New England Biolabs)	(1) New England Biolabs (https://international.neb.com/products/m0380-warmstart-rtx-reverse-transcriptase) (2) Chung, M.T. <i>et al.</i> (2019). Single-cell RT-LAMP mRNA detection by integrated droplet sorting and merging. <i>Lab on a Chip</i> 19, 2425-2434 (3) Garafutdinov, R.R. <i>et al.</i> (2020). The influence of reaction conditions on DNA multimerization during isothermal amplification with Bst exo- DNA polymerase. <i>Appl. Biochem. Biotechnol.</i> 190, 758-771

All given websites and links were accessible on May 18, 2020.