UNDERSTANDING qRT PCR: BASIC PRINCIPLES AND APPLICATIONS

Abstract

The quantitative Reverse Transcription Polymerase Chain Reaction, abbreviated as qRT PCR is a major advancement of the Polymerase Chain Reaction (PCR) technology. Of all the methods available, qRT PCR is the most rapid, sensitive and accurate methodology for measurement of RNA. From the discovery of the enzyme - reverse transcriptase (RT) which led to the development of RT coupled PCR and the major driving force behind identification of the group of RNA viruses, to becoming the most popular diagnostic technique for detection of RNA viruses including SARS CoV-2, the causative agent behind the recent COVID19 pandemic, this chapter describes the technical nuances involved in the conceptualization, refinement, experimental workflow, result interpretation and subsequent wide range of applications in molecular biology, medicine and clinical diagnostics.

Keywords: Polymerase Chain Reaction (PCR) technology.

Author

Dr. Shruti Chowdhari
PhD (CSIR-IGIB)
PDF (IIT Delhi)
New Delhi
India.
shrutichowdhari@gmail.com

I. INTRODUCTION

The discovery of the enzyme **RNA dependent DNA polymerase**, renamed as **reverse transcriptase** led to the conceptualization of "**reverse transcription coupled polymerase chain reaction technique",** also known as **RT PCR.** The classical *in vitro* PCR technique, extensively used for enzymatic amplification (or synthesis) of DNA, cannot be performed directly with RNA, there is need for incorporation of reverse transcription step. The **reverse transcriptase** catalyses the conversion of RNA to complementary DNA (cDNA) followed by amplification of the cDNA with help of PCR. The quantitative analysis of RNA paves way for understanding RNA biology at diverse levels. The RT PCR technique is not only sensitive (i.e. it is able to detect trace quantities of RNA present in a living cell) but also very specific (i.e., it detects the target RNA of interest in a sequence specific manner), thus making RT PCR the cornerstone of every molecular biology laboratory working at RNA level. The qRT PCR method of analysis has nearly replaced the conventional Northen blot analysis technique. More recently, the RT PCR gained widespread popularity all over the world for diagnosis of SARS CoV-2 virus in clinical samples of patients manifesting flue like symptoms during the COVID19 pandemic.

After the discovery of DNA as genetic material (in the 18th century), DNA was recognised as the unit of heredity as well as master blueprint of life. It was widely believed that genetic information within biological systems (prokaryotes and eukaryotes) flows unidirectionally, that is: **DNA to RNA to Protein**, classically known as the **Central Dogma of Molecular Biology**, wherein the RNA is an intermediate carrier of information during protein synthesis. The discovery of RT in RNA tumor viruses simultaneously by David Baltimore at Massachusetts Institute of Technology (MIT), Cambridge, USA and Howard Temin at University of Wisconsin (UW), Madison, USA in the year 1976[(1)] gained worldwide recognition. Although the idea of RT catalysed synthesis of DNA from RNA, contradicted the theory of Central Dogma of Molecular Biology, this concept became the major driving force behind the identification and characterization of group of **RNA viruses** (viruses which are able to synthesize DNA from an RNA template) **and retroviruses**.

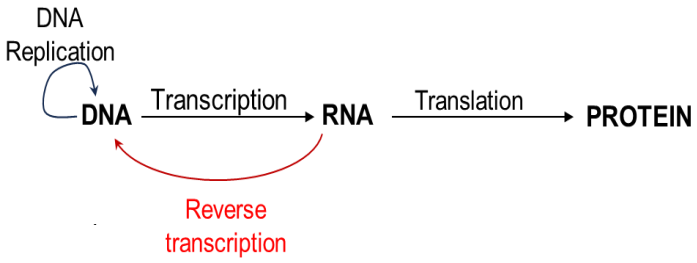


Figure1: Flow of genetic information within biological systems

1. **Basic Principles Along with the Workflow:** The development of quantitative (q) RT PCR not only led to the identification of cancer-causing retroviruses (including human immunodeficiency virus (HIV) and human T cell leukemia viruses (HTLVs)) but also accelerated scientific research involving expression analysis of disease-specific mRNA

biomarkers, profiling several types of RNAs (including ribosomal RNA (rRNA), transfer RNA (tRNA), messenger RNA (mRNA), non-coding RNAs, etc), detection of pathogens (especially RNA viruses) from biological and environmental samples. A typical work flow of the RT PCR set up is described in Fig.2, and basic principles underlining the methodology are described thereafter. A recently developed variation of this technique referred to as digital droplet PCR will be discussed in the applications section.

To begin with, **extraction of RNA** from living cells is the first step and prerequisite to carry out an RT PCR assay for analysis of gene expression or detection of nucleic acids derived from infecting pathogen. In contrast to conventional RNA extraction methods which were labour intensive and prone to contamination at various steps, the method of RNA extraction using the acid guanidinium thiocyanate-phenol-chloroform triphasic separation method described by Chomczynski and Sachhi in the 1987 is most popularly used worldwide. This method is described in the research article entitled, “Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction” in the journal *Analytical Biochem*istry, 1987. Since then, several commercially available ready to use reagents (such as TRIzol (Invitrogen), Qiazol (Qiagen) or Tri reagent (SIGMA), etc) and kits have been commercially developed and are widely available.

The **cDNA synthesis** involves the enzymatic conversion of RNA to complementary DNA (cDNA). The key components of commercially available cDNA synthesis kits include the RT enzyme, its reaction buffer (comprising of Tris-HCL pH 8.4 for maintenance of optimal pH, and enzyme cofactors: KCl, MgCl2), dNTP mixture (comprising of ATP, TTP, GTP, CTP), primers used for reverse transcription (Oligo dT primer or Random hexamer primer), dithiothreitol (DTT), RNase Inhibitor and nuclease free water.

An example of RT used such kits is the RT isolated from **m**olomy **mu**rine **l**eukemia **v**iruses (M-MuLV). The M-MuLV are retroviruses belonging to the gammaretroviral genus of the Retroviridae family. The enzyme possesses RNA-dependent and DNA dependent polymerase activity, but lacks RNase H activity. This is due to a point mutation in the RNase H domain. The point mutation causes a change in the protein structure and function such that the enzyme does not degrade the RNA in RNA-DNA hybrids during synthesis of the first strand cDNA and therefore provides high yields of full-length cDNA even from long RNA templates. The enzyme activity is stable over a wide temperature range (42-55°C) and synthesis of full-length cDNA up to 13 kb is possible. Several genetically modified RTs are being developed which are capable of synthesizing cDNA up to 20 kb.

For cDNA synthesis (i.e. reverse transcription of RNA to cDNA), primers such as Oligo dT primer, random hexamer or gene specific primers can be used.

- The Oligo dT primer, as the name suggests are oligonucleotides (short stretches of nucleic acid residues or polymers) that contain a segment of repeating deoxythymidines (dT). The dT anneal (or bind) to the polyadenosine (polyA) tails of messenger RNA (mRNA) resulting in synthesis of cDNA specific to mRNA fraction of the total RNA extract.

- The random hexamer primers are composed of mixture of single-stranded, random, six-nucleotides (or hexanucleotides) which potentially anneal to random sites in RNA (or DNA) without distinguishing between mRNA or any other species of RNA (like

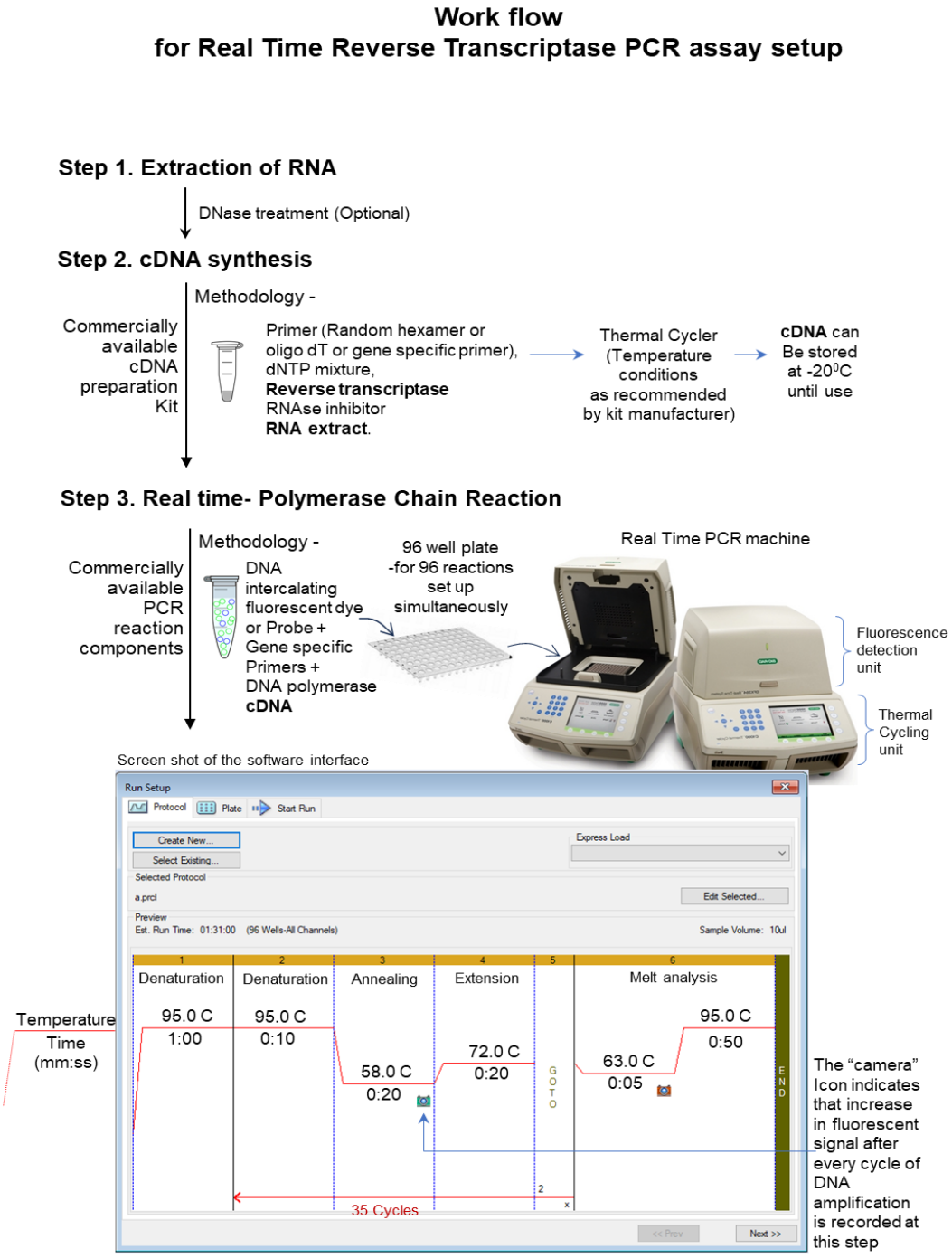


Figure 2: A typical work flow demonstrating setting up of quantitative RT PCR assay ribosomal RNA or RNA originating from virus or bacteria) in the sample and provide a cDNA pool of all types of RNA molecules present in the extract.

- Gene specific primers are used for reverse transcription of RNA corresponding to specific gene of interest which either requires enrichment (as it might be present in low quantity) or specific detection from a total RNA extract.

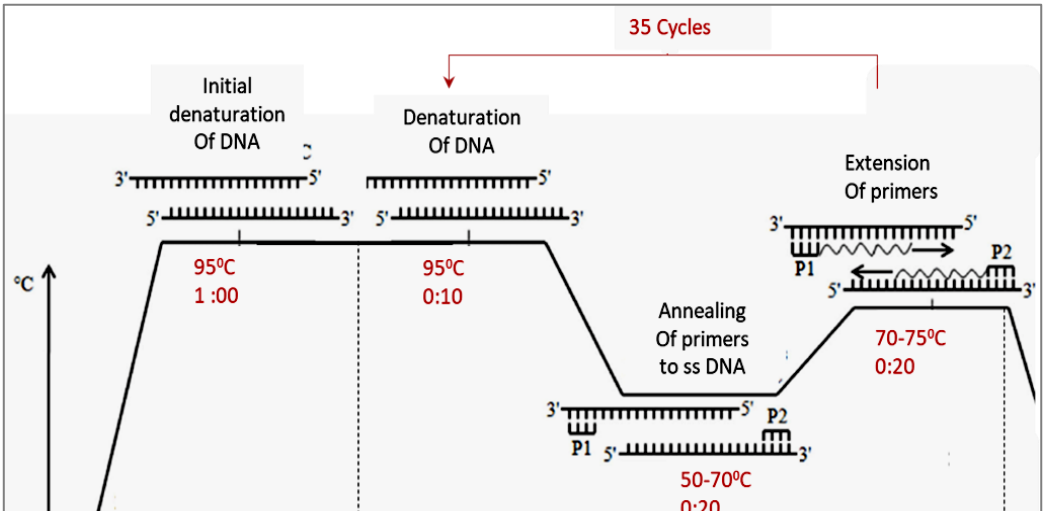


Figure 3a: Steps in a PCR cycle: denaturation, annealing and extension.

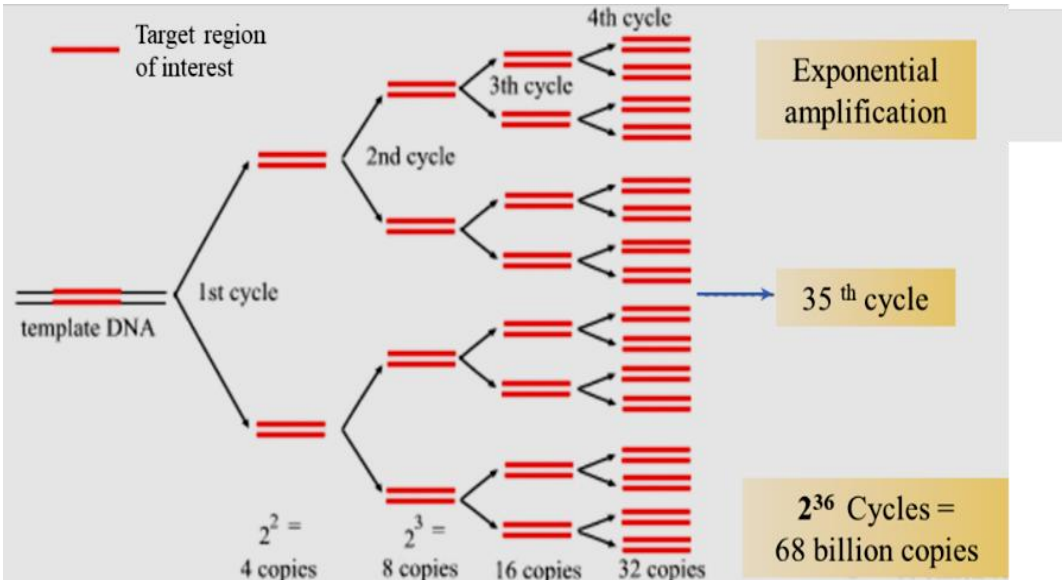


Figure 3b: Exponential amplification of DNA during the PCR cycles

The cDNA can then be subjected to **quantitative PCR amplification** with **gene specific primers** (Forward primer and Reverse primer) using the **real time PCR machine.** The efficiency of the PCR primers contributes significantly to the efficiency of the PCR.

The real time PCR machine comprises of a standard "PCR thermal cycling block" coupled with an "excitation source" (such as a laser or tungsten lamp), "camera" for detection of fluorescence, a computer and software for data analysis. The thermal cycling block

elevates and lowers the temperature of the block in cycles (Fig 3a) in a predefined fashion (the PCR is run for nearly 35-45 cycles) which allows cDNA to be copied exponentially (similar to replication of DNA) as shown in Fig 3b. The camera (or optical detection unit) records the increase in fluorescent signals after each cycle of the DNA amplification. Here, the increase in fluorescent signal is recorded along with each amplification cycle (not at the end of the entire PCR run) hence the acronym "real time". The increase in fluorescent signal indicates the amplification of DNA in each cycle i.e. the fluorescence increases as the number of DNA copies (or amplicons) increase. The real time PCR machine thus enables the detection and analysis of the amplified product without post PCR processing. Most standard PCR machines come with a 96 well block format and allow setting up of the reaction in 96 well reaction plates or PCR strip tubes. For faster performance PCR machines may also come with 384 well block format.

The **idea of "real-time PCR"** which involved the inclusion of a fluorescent reporter system (fluorescent DNA intercalating dye or fluorescent probe) was first described by Higuchi and coworkers in 1992. The continuous monitoring of the accumulation of PCR amplicons in "real-time" leads to certain advantages such as (a) detection of fluorescent signal allows the PCR reaction to proceed in a single closed tube, (b) minimises the risk of carryover contamination, and (c) significantly reduces the time required for completion of PCR assay as it eliminates post-PCR processing when compared to conventional PCR assays. The two most widely used fluorescent reporter systems incorporated for real time analysis include (a) fluorescent DNA-intercalating dyes (such as SyBr Green, Eva Green, etc) or (b) target-specific fluorescently labelled primers and probes. The working principles of both fluorescent reporter systems are described in the diagram below:

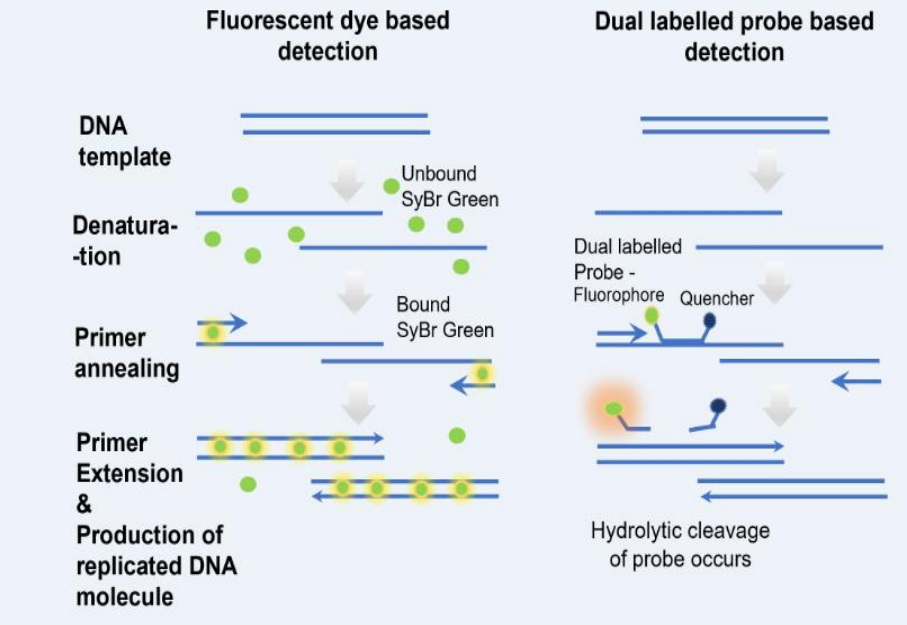


Figure 4: Working principles of the fluorescent reporter systems used in real time PCR. The fluorescence generated by SyBr green bound to dsDNA or by release of the fluorophore after hydrolytic cleavage of the probe is recorded as relative fluorescence units (RFU).

II. DATA ANALYSIS

The kinetics of the PCR amplification are visualized with help of the instrument-software (for example, Biorad CFX Maestro for Biorad qPCR machines). The software allows the visualization of fluorescence generated/cycle (which is directly proportional to increase in DNA amount per cycle) in the form of an **amplification curve**. The fluorescence in terms of relative fluorescence units (RFU) is plotted along the Y axis, whereas the number of PCR cycles is represented along the X axis.

1. **The Amplification Curve:** The increase in fluorescence per cycle is demonstrated by the green line and resembles a sigmoidal pattern of amplification. The amplification curve can be divided in three phases as shown in the fig 5a:

   - **Background (or baseline):** At the beginning of the run (usually between cycles 0 and 15), when the amount of PCR product is low, little fluorescence is emitted and this fluorescence is indistinguishable from the background signals (background signals may occur due to autofluorescence of reaction components). This section of cycles is known as the background signal or noise. The point where signals generated due to PCR amplification crosses the background fluoresce is considered by the instrument as the "threshold".
   - **Exponential phase**: Once the amount of amplified PCR product and the corresponding fluorescence exceeds the background fluorescence level, the reaction enters the exponential phase. Here, the amount of PCR product doubles for every PCR cycle. This can be observed in cycles 15 to 30.

   - **Plateau:** Once all of the reaction components (including nucleotides) have been used up in the PCR reaction, further amplification will slow down and ultimately reach a plateau phase. This is the region where no more PCR products can be synthesized. This can be usually be seen in cycles 30 to 40 of the amplification plot.

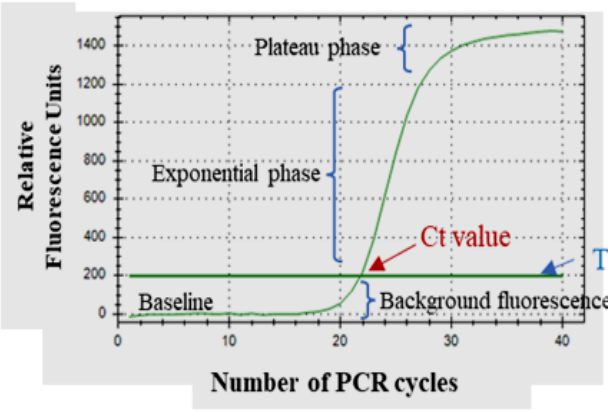
2. **Ct Value:** The Ct value also known as "Cycle threshold" is defined as the number of cycles after which the fluorescence signal of the specific PCR product crosses the threshold background signals. It is also referred to as the Cq or quantification cycle.

   The Ct value and the amount of DNA present in the PCR are inversely correlated. This means that a PCR assay starting with abundant amount of input DNA tends to demonstrate an early Ct value and a PCR assay starting with low amount of input DNA tends to demonstrate late Ct value.

   The real time PCR data can be analysed using several methods. Two widely employed methods of representation of real time PCR data include:

   - Absolute quantitation using standard curve analysis.
   - Relative quantitation using $2^{-\Delta\Delta Ct}$ method.

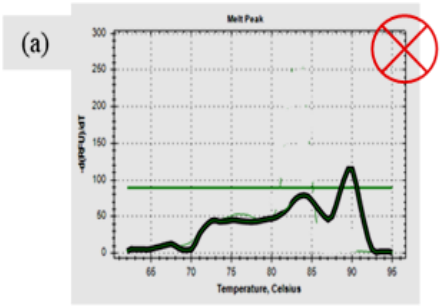
(a) Amplification curve :



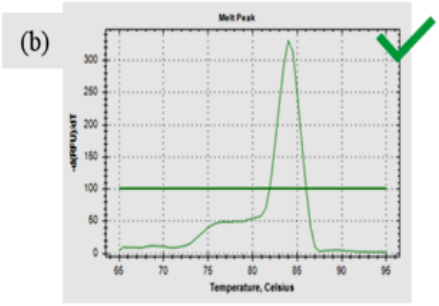
The PCR Cycle number at which the fluorescence generated by each amplified DNA sample crosses the threshold is referred to as the "Ct" value

Threshold set by the instrument

(b) Melting curve : Each amplification curve has a corresponding melting curve

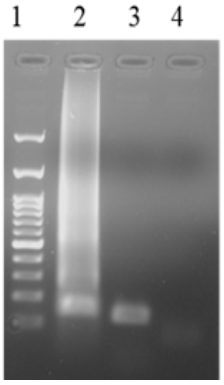


Multiple peaks indicates multiple (non-specific) amplicons in the PCR product.



Single peak indicates single (specific) amplicon in the PCR product

(c) Below is the agarose gel electrophoresis picture of the RT PCR products corresponding to the above mentioned melt peaks (a, b).



1: 100bp DNA ladder

2: The smear of the PCR product corresponds to the multiple peaks as shown in (a), this indicates non specific amplification because of inefficient primers.

3: The sharp band of the PCR product corresponds to the single sharp peak at 84^0C as shown in (b), this indicates specific amplification by primers.

4: No amplification, only primer dimer visible, this is usually represented by a small peak near 75^0C in the melting curve.

Figure 5: Visualization of the real time PCR kinetics.

3. **Absolute quantitation using standard curve analysis:** This method is used to determine the precise quantity of DNA or RNA present in the biological sample. In this method, first the standard curve is prepared by plotting the Ct values obtained from the qPCR for serially diluted copies of DNA of known concentration.

   Usually, the standard curve is a straight line and is represented by the mathematical equation: y = mx + c, wherein y is the Ct value, m is the slope of the straight line and x describes the concentration of DNA, c is the y intercept value. Subsequently, the Ct value obtained for the test sample (of unknown concentration) can be evaluated using the equation (y = mx + c) of the standard curve prepared for the particular gene being analysed.

   For example, this method is often used for calculating viral load (or microbial load) in biological specimens. The evaluation of microbial load also uses the assumption that single cell of a microorganism (or virus) comprises of one copy of its whole genome and the genome comprises of only one copy of each gene. Thus counting the number of copies of the gene of interest amplified via qPCR provides information about the number of microbial cells (or viral load in case of viral infection) present in a biological sample of an infected individual. The slope (m) of the standard curve provides an idea about efficiency of the PCR. A disadvantage of absolute quantification includes the increased effort to generate standard curves.

4. **Relative quantitation using $2^{-\Delta\Delta Ct}$:** It is also known as comparative quantitation. Here the change in levels of mRNA transcribed by gene of interest relative to mRNA transcribed by a second gene called the internal control or calibrator gene is measured. The internal control gene is usually a gene which abundantly expressed in cells and its expression level is unaffected by the condition under investigation. The choice of internal control gene depends on several factors such as cell type being studied, stability of the mRNA in a particular condition, or species under investigation. For example, this method is widely employed during quantitative analysis of gene expression in given conditions and given timepoint.

   After the qPCR run is complete, the data obtained is obtained in the form of Ct values in addition to few other reaction kinetics related parameters. The change in expression of gene of interest is calculated using the formula $\mathbf{2^{-\Delta\Delta Ct}}$ as described below:

   ΔCt = Ct (Gene of interest) – Ct (internal control).

   $\Delta\Delta Ct$ = ΔCt (Treated) – ΔCt (Un-treated/control).

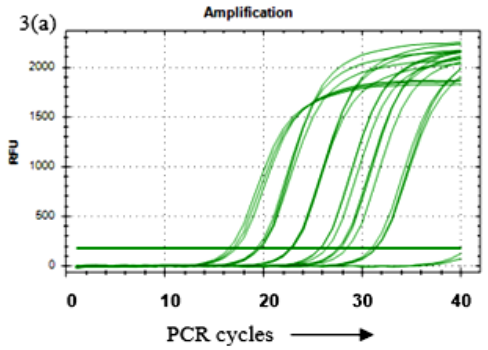
   Fold change in gene of interest in treated (any context) versus untreated (control) = $\mathbf{2^{-\Delta\Delta Ct}}$

   Examples of internal control genes include GAPDH, β-actin, 18s rRNA, etc for eukaryotic organism, 16s rRNA for prokaryotic organism, RdRp for certain viruses.

The RT PCR tests are referred to as “quantitative” as this method allows quantification of the specific DNA present in a sample.

Absolute quantitation :

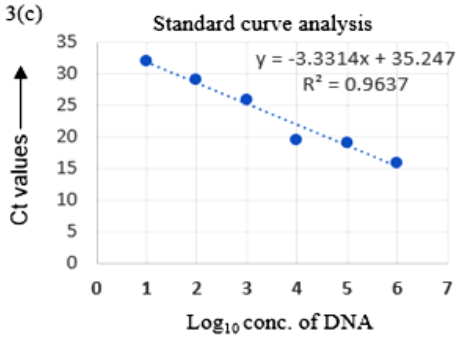
3(a)



3(b)

Conc. of DNA (calibrator)	Log $_{10}$ Conc.	Ct (mean of 3 replicates)
0	0	0
10	1	32.01
100	2	29.02
1000	3	25.99
10000	4	19.50
100000	5	19.00
1000000	6	16.00

3(c)



y= mx+c (equation of the line)

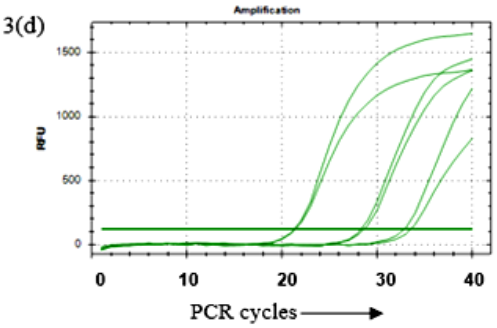
y = -3.3314x +35.247

Here: Y= Ct Value; x = Log_{10} (Viral load)

Gene copies/cell = 10^x

Relative quantitation:

3(d)



For example:
Gene of interest (GOI) is TP53
Internal control (IC): 18s rRNA

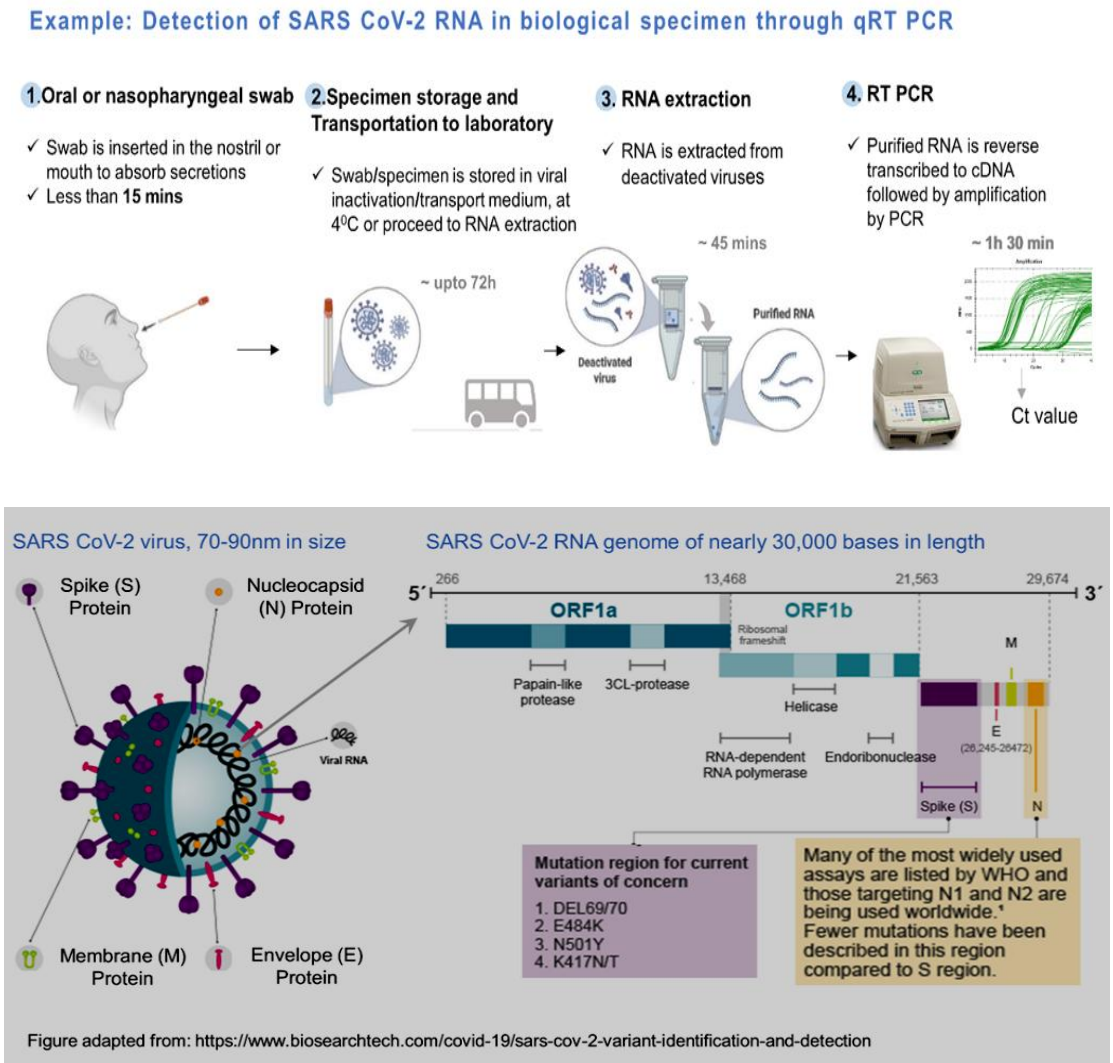
	Ct (GOI: TP53)	Ct (IC: 18s rRNA)	Δ Ct: Ct(GOI) - Ct(IC)	ΔΔ Ct: Δ Ct (Treated)- Δ Ct (Untreated)	Fold change $2^{-\Delta\Delta Ct}$
Cells (UV treated)	29.1	21.5	7.6		
				-3.3	9.84
Cells (Untreated)	32.2	21.3	10.9		

Inference: The expression of TP53 is 9.8 ford higher in UV treated cells compared to untreated cells.

Figure 6: Calculations for analysis of data using absolute and relative quantification method

III. APPLICATIONS

1. **Molecular diagnostics:** Quantitative reverse transcriptase PCR (qRT-PCR) is now being recognised globally as an important diagnostic tool for nucleic acid amplification-based detection of pathogens especially RNA viruses in biological specimens. Until now, this technique has been successfully employed for detection of clinically prevalent RNA viruses including enteroviruses, west-nile virus, dengue virus, human metapneumovirus, hantavirus, and more recently SARS CoV-2 virus, the causative agent of COVID19 pandemic. Among the few changes forced by the recent SARS CoV-2 pandemic (2020-22) the installation of real time PCR machines in molecular diagnostic facilities across the world has led the medical fraternity to turn to RT-PCR based diagnostics of infectious disease causing pathogens above culture-based assays. More recently, single tube RT PCR assays (1-step RT PCR) have also been developed which include sequential conversion of RNA to cDNA (reverse transcription) followed by cDNA amplification in a single tube format thereby surpassing the need for transfer of material from one tube to another, thereby surpassing the risk of spillage and cross-contamination of the infectious disease specimen.



2. **Basic biological sciences research:** As RNA is an important intermediate in the flow of information during gene expression (DNA to RNA to PROTEIN), the RNA is quantified with help of RT PCR and this data is used to understand aspects such as:

    - Difference in expression of a gene in the diseased state compared to the normal state.
    - Experimental validation of the extent of transcription of a gene.
    - Gene expression changes occurring in cells exposed to a chemical substance (e.g., drug, toxin, hormone or cytokine).
    - Gene expression changes during cell differentiation or development.
    - To validate the effectiveness of small interfering RNA or antisense oligonucleotides.
    - Quantification of expression of noncoding RNA.

3. qRT-PCR assays are also frequently used for the detection of **viral load while** monitoring the response of the patient to therapy. For example, after appropriate antiretroviral therapy, the HIV positive patient must demonstrate an elevated CD4 T cell count and a reduction in HIV viral load.

4. This method may also be used for detection of live versus dead microorganisms such as bacteria, fungi and parasites by targeting their RNA, as ratio of RNA to DNA of any microorganism in any kind of sample indicates the presence of **viable** organisms.

5. qRT PCR is increasingly being used in environmental microbial ecology research as a tool for detection and quantification of microbes enriched in wastewater microbial ecosystem. Development of qRT PCR based evaluation of biomarkers of fecal contamination (including enteric RNA viruses and bacteriophages) in treated wastewater before reuse is an area of active research.

6. Incorporation of RT PCR assay in "Point of Care" diagnostics is an area or ongoing research.

7. Digital droplet PCR: Over the years, the PCR technology has undergone tremendous modifications broadly classified into three generations. The first generation of PCR depends on post PCR gel electrophoresis for evaluation of PCR amplified products. It is limited by poor detection limit, is cumbersome, and is useful only for qualitative analysis. The second generation of PCR, which is the quantitative real time PCR (qRT PCR), can quantify the products with help of standard curves, but also show low tolerance to interfering substances. The third generation: recently developed droplet-based digital PCR (ddPCR) is based on partitioning of the sample into thousands of nanoliter-sized droplets (~ 20,000 droplets) such that each droplet must contain 0 or 1 DNA molecule. Each droplet undergoes PCR amplification of the target sequence. Thereafter, the fluorescent signal of each droplet is recorded with help of an automated droplet reader. If the droplet contains at least one copy of the target sequence, it is a positive droplet (fluorescent droplet). If the droplet does not contain any target copies, it is a negative droplet (exhibiting little to no fluorescence). This ratio of positive droplets to negative droplets is then evaluated with Poisson statistics to enumerate the concentration of the DNA template in the original sample. This technique has demonstrated high sensitivity, accuracy, and reproducibility. This technology is taken its place in clinical settings for

applications such as trace DNA detection, rare mutation detection, copy number variation gene expression analysis and many more[6,7].

REFERENCES

[1] Annu Rev Virol. 2016 Sep 29;3(1):29-51. doi: 10.1146/annurev-virology-110615-035556. Epub 2016 Jul 22. PMID: 27482900
[2] Verma, I.M., Reverse transcriptase, The Enzymes (Boyer, P.D., ed), Academic Press Inc., vol. 14, 87-103, 1981.
[3] Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem. 1987 Apr;162(1):156-9. doi: 10.1006/abio.1987.9999. PMID: 2440339.
[4] Gerard, G.F. and D'Alessio, J.M., Methods in Molecular Biology, 16, Humana Press, Totowa, N.J., 73-93, 1993.
[5] Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc. 2008;3(6):1101-8. doi: 10.1038/nprot.2008.73. PMID: 18546601.
[6] Zhu H, Zhang H, Xu Y, Laššáková S, Korabečná M, Neužil P. PCR past, present and future. Biotechniques. 2020 Oct;69(4):317-325. doi: 10.2144/btn-2020-0057. Epub 2020 Aug 20. PMID: 32815744; PMCID: PMC7439763.
[7] Heinzelmann D, Lindner B, Renner B, Fischer S, Schulz P, Schmidt M. Droplet digital PCR: A comprehensive tool for genetic analysis and prediction of bispecific antibody assembly during cell line development. N Biotechnol. 2023 Dec 25;78:42-51. doi: 10.1016/j.nbt.2023.10.001. Epub 2023 Oct 3. PMID: 37797917.

