Chapter 4

Detection of Rubella Virus by Tri-Primer RT-PCR Assay and Genotyping by Fragment RT-PCR

Suji George

Abstract

Polymerase chain reaction (PCR) is a widely used technique in the diagnosis of viral infections due to its low cost, high sensitivity, and specificity. Although the more advanced variations of PCR, such as real-time PCR and digital PCR are now available to researchers, conventional PCR is still used in many research studies. Here we describe the protocol for tri-primer diagnostic reverse transcription polymerase chain reaction for detection of rubella in throat swabs and further detailed protocol for a two fragment genotyping using two different sets of primers. In tri-primer diagnostic PCR, one forward and two reverse primers are used to detect clade I and clade II of the rubella virus. In the two fragments genotyping, each fragment of the genome is amplified, sequenced separately, and then the overlapping regions are aligned and full length sequence window is obtained.

Key words Rubella virus, Congenital rubella syndrome, Genotyping, RT-PCR, Surveillance

1 Introduction

The *Rubella virus* belongs to the family *Matonaviridae*, specifically to the genus *Rubivirus* within it. Rubella virus was first distinguished from measles and scarlet fever, both exanthematous diseases, in 1814, by George de Maton. It was in honor of him that a viral family was established, called *Matonaviridae* [1]. The Rubella virus was initially classified into the *Togaviridae* family, which literally stood for "shrouded" viruses, owing to their envelope [2]. Members of the *Togaviridae* family include the *Alphavirus* genus [2], of which there are some well-known arthropod-borne viruses (arboviruses), like *Chikungunya virus, Sindbis virus, Eastern Equine Encephalitis virus, Western Equine Encephalitis virus, Venezuelan Equine Encephalitis virus, Semliki Forest virus, Eilat virus*, among many others. What differentiates the *Rubella virus* from the above-mentioned members of the *Alphavirus* genus is its mode of transmission being respiratory droplet-mediated instead of by arthropod vectors [3]. The entire genus of *Rubivirus* had its family

Chhandak Basu (ed.), *PCR Primer Design*, Methods in Molecular Biology, vol. 2392, https://doi.org/10.1007/978-1-0716-1799-1_4,

changed from the *Togaviridae* to the *Matonaviridae* [4]. As of now, there are only three members of the matonaviridae family- *Rubella virus, Ruhugu virus, and Rustrela virus* [4]. The latter two viruses were shown to be considerably capable of crossing species barriers, and so it was eventually speculated that the *Rubella virus* may also have been originally of zoonotic origin.

There is a vast public health concern regarding the *Rubella virus.* Historically, the virus and its disease have been associated with populations such as the military, boarding schools, and other confined populations [5]. The National Communicable Disease Center (now called the Centers for Disease Control and Prevention (USA)) had, since then, established surveillance networks for the monitoring of rubella. This was done to survey the need for vaccination by keeping in mind the epidemiological parameters of the disease throughout the United States. Since the introduction of rubella vaccination in 1969, progress led to the declaration of elimination of rubella endemicity in the USA by 2005 [6], accounted for by <25 cases per year since 2001, 95% and above vaccination coverage in school children and an estimated 91% of the population is immune to the virus. Two other important aspects leading to this achievement were adequate surveillance to detect rubella outbreaks and the genotyping of viruses that showed their origin from other parts of the world. The diagnosis and surveillance of disease are indeed incomplete without the modern advancements of molecular techniques that have empowered the epidemiologists and the diagnosing personnel. The Polymerase Chain Reaction (PCR), developed by Kary Mullis, has evolved by leaps and bounds to open up brand new avenues in biology altogether, and yet its more basic form remains a mainstay of clinical diagnosis and research to this day [7].

The viral genome is of a single-stranded positive sense Ribonucleic Acid (RNA) [8]. It is approximate 10 kilobases length, composed of the following genes—P150, P90, C (capsid), E1, and E2—from the 5′ to the 3′ end. P150 and P90 comprise the nonstructural protein Open Reading Frame (ORF) proximal to the 5′ end of the genome (Fig. 1). These two proteins have a function in the replication of viral RNA. The remaining three are the structural proteins, encoded from the structural protein ORF. The C protein encodes for the core of the virion, whereas E1 and E2 encode for the envelope proteins of the virion. The structural ORFs are, however, translated from a subgenomic stretch of RNA synthesized from the viral genome containing only the C, E1, and E2 genes [9]. There are Untranslated Regions (UTRs) at the 5′ and 3′ ends as well as junction regions between the ORFs. The World Health Organization has developed a taxonomy system of Rubella viruses which make use of a 739 nucleotides sequence (nucleotides 8731-9469) within the E1 gene. Additionally, the structural protein ORF of some viruses is also used for classification. The Rubella

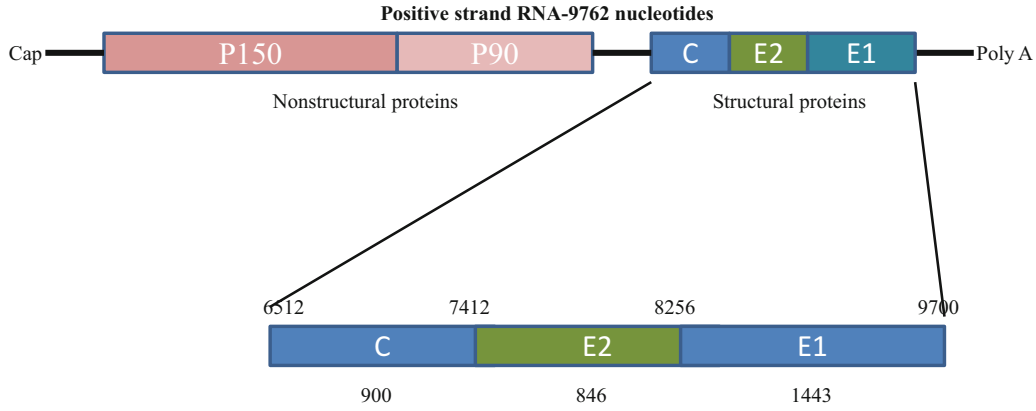

Fig. 1 Positive strand genome of rubella virus

viruses are serologically monotypic. Thus, two clades corresponding to genotype I and genotype II were established, with clade I having ten genotypes—1a, 1B, 1C, 1D, 1E, 1F, 1G, 1h, 1i, and 1j—while clade II has three genotypes—2A, 2B, and 2C [10]. Lowercase letters indicate provisional genotypes, while uppercase ones indicate recognized genotypes. The maximum divergence among clade I viruses is 5.8%, while that among clade II viruses is 8.0%. The inter clade divergence is about 8.2% [11]. Clade I viruses have a worldwide circulation whereas clade II viruses are more restricted to Eurasia. RV virus has a high GC content of 69.6% for the whole genome, 73% for the C coding region, 71% for the E2 coding region, and 66.5% for the E1 coding region. The high GC content often leads to the formation of a stable inherent secondary structure that can either inhibit reverse transcriptase or can affect primer annealing. The recommended sequence window for genotyping is 739 bp in the E1 coding region and taking into account primer binding sites and the poorly resolved peaks of sequence near the primer binding region, the RT-PCR amplicon needs to be much larger (e.g., 900 bp). Thus larger the amplicon, the higher is the minimum copy number required for detection that means a lower sensitivity. In order to circumvent the issue of sensitivity, a two fragment system RT-PCR has been established for the genotyping of rubella. Rather than amplifying a single larger amplicon of 900 bp, two overlapping amplicons of size 480 bp and 633 bp are amplified and the sequence is then merged to obtain the 739 window region.

In this chapter, a detailed RT-PCR protocol is described using three primers, one forward primer common to both the clades and two separate reverse primers to detect both clade I and clade II genotype [12]. Further, the detailed protocol for genotyping is discussed using two different polymerase enzyme systems

SuperScript™ III One-Step RT-PCR System and Qiagen one step RT-PCR Kit using two sets of primers.

2 Materials

2.1 RNA Preparations

1. QIAamp Viral RNA Isolation Kit (Qiagen).
2. 70% ethanol.
3. Sterile 1.5 ml microcentrifuge tubes.
4. Pipette tips with aerosol barriers.

2.2 RT-PCR

SuperScript™ III One-Step RT-PCR System with Platinum™ *Taq* DNA Polymerase ((Invitrogen #12574-018)).

Betaine, 5 M (Sigma #B0300-1VL) if SS III enzyme kit is used (*see* **Note 1**).

QIAGEN OneStep RT-PCR Kit ((Qiagen CAT# 210210 or 210212)).

2.3 Primers for Diagnostic PCR

Diagnostic PCR gives information regarding the presence of rubella virus in body fluid. The diagnostic PCR for rubella amplifies a region of 185 base pair in the E1 coding region of the genome (Fig. 2).

Forward Primer: 5′-CAA CAC GCC GCA CGG ACA AC-3′.

Reverse Primer 1: 5′-CCA CAA GCC GCG AGC AGT CA-3′.

Reverse Primer 2: 5′-CCA CGA GCC GCG AAC AGT CG-3′.

2.4 Primers for Genotyping

Genotyping is important to characterize the circulating virus strain as it gives information regarding prevalence of the virus in the population and also provides crucial information regarding success of vaccination campaign. Further continued genotyping in surveillance studies are also important in confirming the disappearance of endemic strains at the eliminating or eliminated stage.

2.4.1 Primer Pair-1

Primer pair-1 amplifies fragment-1 (480 bp region) in the E1 gene of the rubella genome (Fig. 3).

Forward Primer1-1: 5′-AGCGACGCGGCCTGCTGGGG -3′.

Reverse Primer1-1: 5′-GCGCGCCTGAGAGCCTATGAC -3′.

2.4.2 Primer Pair-2

Primer pair-2 amplifies fragment-2 (633 bp region) in the E1 gene of the rubella genome (Fig. 3).

Forward Primer2-2: 5′-TGGGCCTCCCCGGTTTG -3′.

Reverse Primer2-2: 5′-CGCCCAGGTCTGCCGGGTCTC -3′.

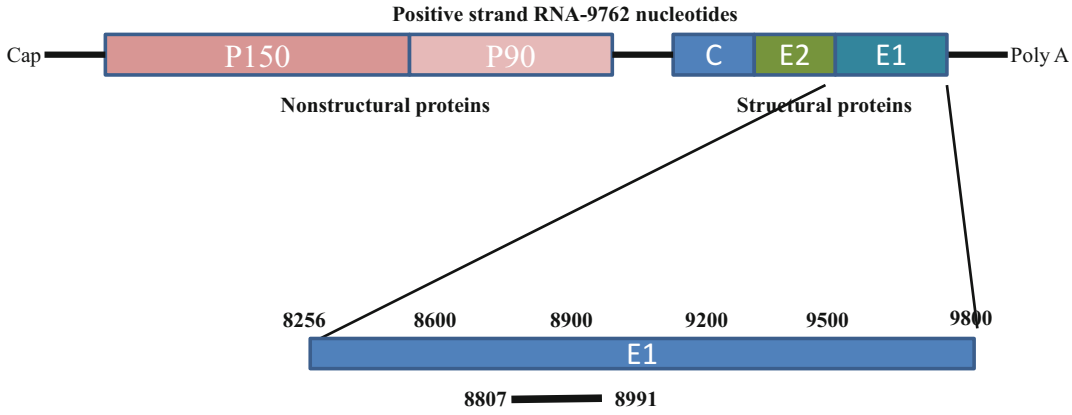

Fig. 2 The region between 8807 and 8991 is amplified during diagnostic PCR. The scale shown in the figure is arbitrary and is used for representing the amplified region during diagnostic PCR

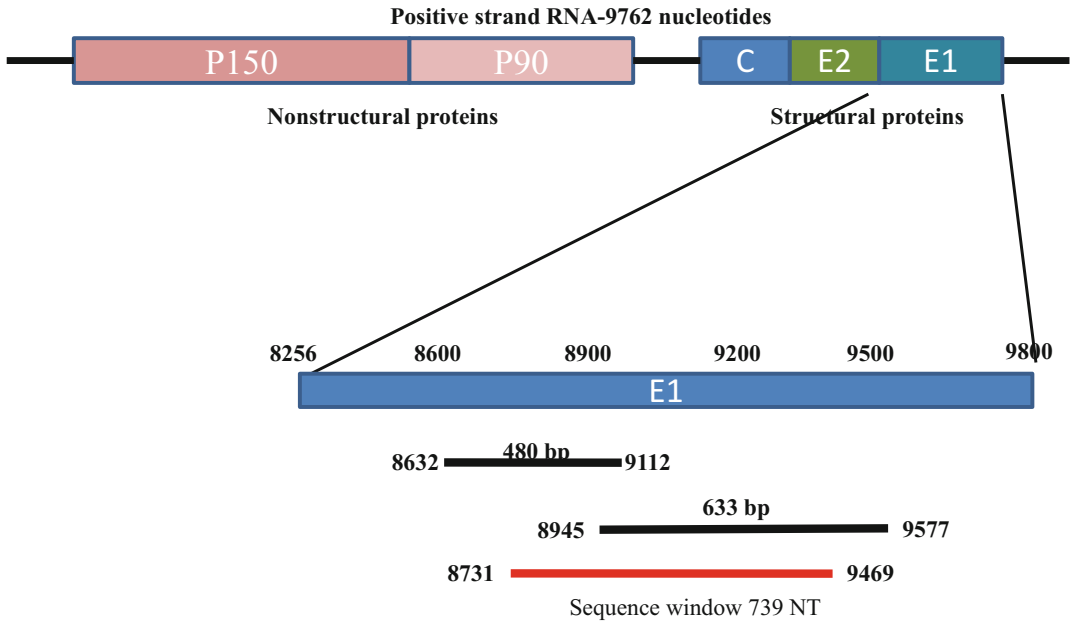

Fig. 3 In two fragment rubella, genotyping fragment-1 (region between 8632 and 9112) and fragment-2 (region between 8945 and 9577) are amplified separately and the resulting sequence is aligned and merged to obtain the sequence window of 739 base pairs (region between 8731 and 9469). The scale shown in the figure is arbitrary and is used for representing the amplified region during genotype PCR

2.5 Gel Electrophoresis

Agarose.
Gel loading dye 6×.
Invitrogen™ SYBR™ Safe Stain.

2.6 Sanger Sequencing

PCR Purification Kit (Qiagen).
Ethylenediaminetetraacetic Acid (125 mM).

0.698 g in 15 ml of Milli q water.

Sodium acetate (3 M).

3.69 g in 15 ml of Milli q water.

Hi-Di Formamide (Applied Biosystems™).

BigDye Terminator version 3.1 cycle sequencing kit (Applied Biosystems™).

2.7 Sequence Analysis

MEGA, Staden (DNA sequence assembly, editing software).

3 Methods

3.1 Clinical Specimen

Throat swab samples collected from patients with clinically suspected rubella virus infection.

3.2 RNA Preparations

The protocol for the extraction of RNA is common for both diagnostic PCR and genotype PCR.

Viral RNA from throat swabs was extracted using the following protocol:

1. **The QIAamp^{R} viral RNA kit includes**: Buffer AVL, Buffer AVE, Buffer AW1, and Buffer AW2, Carrier RNA (Poly A), QIAamp Mini Spin Columns, and collection tubes.
2. **Dilution of carrier RNA**: 310 µl of Buffer AVE is added to 310 µg of lyophilized carrier RNA to obtain a solution of 1 µg/ml. Aliquot the carrier RNA solution and store it at −30 °C to −15 °C.
3. **Preparation of Buffer AVL-Carrier RNA**: Add 5.6 µl of carrier RNA-AVE mix to 0.56 ml Buffer AVL. This amounts to 5.6 µg of per sample carrier RNA (carrier RNA as lyophilized powder is constituent of the The QIAampR viral RNA kit). For larger numbers of sample, the volume can be calculated by the following formulae:

$$n \times 0.56 \text{ ml} = y \text{ ml}$$

$$y \text{ ml} \times 10 \text{ µl/ml} = z \text{ µl}$$

where: n = number of samples to be processed simultaneously.

y = calculated volume of Buffer AVL.

z = volume of carrier RNA–Buffer AVE to add to Buffer AVL.

4. Pipette 560 µl of buffer AVL-carrier RNA mix in 1.5 ml of microcentrifuge tube. Add 140 µl of throat swab samples to it and mix by pulse vortexing for 15 s. Incubate at room

temperature for 10 min and briefly centrifuge the tube to remove the drops from inside the lid.

5. Add 560 µl of ethanol (96–100%) to the sample and then mix by pulse vortex for 15 s. Then centrifuge the tube to remove the drops from inside the lid.
6. Load 630 µl of solution from **step 5** to the QIAamp mini column (in 2 ml of collection tube). Centrifuge at 6000 × *g* (8000 rpm) for 1 min. Transfer the QIAamp mini column into a clean 2 ml collection tube and discard the tube containing the filtrate.
7. Carefully open the QIAamp mini column and repeat **step 6**.
8. Wash the QIAamp mini column with 500 µl buffer AW1. Centrifuge at 6000 × *g* (8000 rpm) for 1 min. Place the QIAamp Mini column in a clean 2 ml collection tube, and discard the tube containing the filtrate.
9. Wash the QIAamp mini column with 500 µl buffer AW2. Centrifuge (20,000 × *g*; 14,000 rpm) for 3 min. Place the QIAamp Mini column in a clean 2 ml collection tube, and discard the tube containing the filtrate.
10. Centrifuge the QIAmp mini column (In new 2 ml of collection tube) at full speed for 1 min.
11. Transfer the QIAmp mini column to 1.5 ml of centrifuge tube. Add 50 µl of buffer AVE equilibrated to room temperature and incubate at room temperature for 1 min.
12. Centrifuge at 6000 × *g* (8000 rpm) for 1 min. The eluted viral RNA can be stored at −80 °C.

3.3 Assay Controls

Negative and positive controls must be run in each assay.

3.3.1 Positive Controls

Positive control for diagnostic PCR is RNA transcript of the rubella E1 region of the vaccine strain RA 27/3.

3.3.2 Negative Controls

Instead of adding RNA, an equivalent volume of only nuclease-free water is added.

Water Controls

Extraction Controls

During RNA extraction, mock extraction control consisting of only nuclease-free water is used.

3.4 For Diagnostic PCR

The Viral RNA extracted using the above protocol is used as template for the diagnostic PCR and the protocol described below uses the SuperScript™ III One-Step RT-PCR System.

1. **Thaw the following on ice**: SuperScript™ III One-Step RT-PCR System with Platinum™ *Taq* DNA Polymerase kit

Table 1
RT-PCR Mix for Rubella diagnostic PCR

Components	Volume (μl)
2× Reaction Mix	12.5
Forward primer (10 μM)	0.5 (0.2 μM)
Reverse primer 1 (10 μM)	0.5 (0.2 μM)
Reverse primer 2 (10 μM)	0.5 (0.2 μM)
SuperScript™ III RT/Platinum™ Taq Mix	1.0
Template RNA	4.0
RNase-free water	6.0
Total	25

components, Forward Primer, Reverse Primer 1, and Reverse Primer 2.

2. Prepare PCR mix as per Table 1 and chill on ice.
3. Perform the reverse transcript step for 30 min at 50 °C and denaturation for 5 min at 95 °C followed by 35 cycles of 95 °C for 30 s, 60 °C for 30 s, and 72 °C for 1 min followed by 72 °C for 10 min to amplify a 185-nucleotide region in the E1 Coding region (Nucleotides 8807 to 8991).
4. Load 10 μl of the PCR product after adding appropriate amounts of loading dye to the wells of 2% agarose gel. Run electrophoresis in TAE at 120 V for 45 min until the bands resolve.

3.4.1 Rubella Genotyping

The viral RNA extracted in the above step is used as template for rubella genotyping. Described here are rubella genotyping using two enzyme systems [1] SuperScript™ III One-Step RT-PCR System and [2] Qiagen One-Step RT-PCR System kit.

3.4.2 Rubella Genotyping (Using the Superscript III Kit)

1. **Thaw the following on ice**: SuperScript™ III One-Step RT-PCR System with Platinum™ *Taq* DNA Polymerase kit components, Forward Primer1-1, Reverse Primer1-1, Forward Primer2-2, and Reverse Primer2-2.
2. Prepare PCR mix as per Tables 2 and 3 and chill on ice. Here two separate PCRs one with Primer pair-1 and the other with Primer pair-2 were set up.
3. Perform the reverse transcript step for RT step for 30 min at 55 °C and denaturation for 2 min at 94 °C. Further the reaction mixture was incubated for 40 cycles of 94 °C for 15 s, 55 °C for 30 s, and 68 °C for 1 min followed by 68 °C for 5 min for superscript III.

Table 2
RT-PCR Mix for Rubella genotype PCR using the primer pair-1 and Superscript III enzyme

Components	Volume (μl)
2× Reaction Mix	25.0
Betaine 5 M	10.0 (1 M)
Forward primer1-1 (10 μM)	1.0 (0.2 μM)
Reverse primer1-1 (10 μM)	1.0 (0.2 μM)
SuperScript™ III RT/Platinum™ Taq Mix	2.0
Template RNA	4.0
RNase-free water	17.0
Total	50

Table 3
RT-PCR Mix for Rubella genotype PCR using the primer pair-2 and Superscript III Kit

Components	Volume (μl)
2× Reaction Mix	25.0
Betaine (5 M)	10.0 (1 M)
Forward primer2-2 (10 μM)	1.0 (0.2 μM)
Reverse primer2-2 (10 μM)	1.0 (0.2 μM)
SuperScript™ III RT/Platinum™ Taq Mix	2.0
Template RNA	4.0
RNase-free water	7.0
Total	50

4. Load 40 μl of the PCR product after adding appropriate amounts of loading dye to the wells of 1.5% agarose gel. Run electrophoresis in TAE at 120 V for 45 min until the bands resolve and excise the band. The PCR products were then purified using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA, USA). For sequencing, the amplicons were first subjected to sequencing PCR using the dye terminator method (BigDye Terminator version 3.1 cycle sequencing kit; Applied Biosystems, Foster City, CA) and followed by precipitation by EDTA/NaOAc/Ethanol. The precipitate was dissolved in Hi-Di Formamide and bidirectional sequence reading was done using the ABI 3100 Genetic Analyzer (Applied Biosystems, Hitachi, Tokyo ABI PRISM 3100). Sequence data were aligned, edited, and assembled to obtain the 739-nt sequence window.

Table 4
RT-PCR Mix for Rubella genotype PCR using the primer pair-1 and Qiagen Kit

Components	Volume (μl)
5× QIAGEN OneStep RT-PCR Buffer	10.0
dNTP Mix	2.0
5× Q-solution	10.0
Forward primer1-1 (10 μM)	2.0 (0.2 μM)
Reverse primer1-1 (10 μM)	2.0 (0.2 μM)
QIAGEN One Step RT-PCR Enzyme Mix	2.0
Template RNA	4.0
RNase-free water	18.0
Total	50

3.4.3 Rubella Genotyping (Using the Qiagen One Step RT-PCR Kit)

1. **Thaw the following on ice**: Qiagen One-Step RT-PCR System kit components, Forward Primer1-1, Reverse Primer1-1, Forward Primer2-2, and Reverse Primer2-2.
2. Prepare PCR mix as per Tables 4 and 5 and chill on ice. Here two separate PCRs one with Primer pair-1 and the other with Primer pair-2 were set up.
3. Cycling condition for Qiagen one step RT-PCR kit (Qiagen, Hilden, Germany) involved RT step for 30 min at 50 °C and denaturation for 15 min at 95 °C. Further the reaction mixture was incubated for 40 cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 1 min followed by 72 °C for 10 min. Load 40 μl of the PCR product after adding appropriate amounts of loading dye to the wells of 1.5% agarose gel. Run electrophoresis in TAE at 120 V for 45 min until the bands resolve and excise the band. The PCR products were purified using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA, USA). For sequencing, the amplicons were first subjected to sequencing PCR using the dye terminator method (BigDye Terminator version 3.1 cycle sequencing kit; Applied Biosystems, Foster City, CA) and followed by precipitation by EDTA/NaOAc/Ethanol. The precipitate was dissolved in Hi-Di Formamide and bidirectional sequence reading was done using the ABI 3100 Genetic Analyzer (Applied Biosystems, Hitachi, Tokyo ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, Hitachi, Japan). Sequence data were aligned, edited, and assembled using MEGA [13] and Staden [14] to obtain the 739-nt sequence window.

Table 5
RT-PCR Mix for Rubella genotype PCR using the primer pair-2 and Qiagen Kit

Components	Volume (μl)
5× QIAGEN OneStep RT-PCR Buffer	10.0
dNTP mix	2.0
5× Q-solution	10.0
Forward primer2-2 (10 μM)	2.0 (0.2 μM)
Reverse primer2-2 (10 μM)	2.0 (0.2 μM)
QIAGEN One Step RT-PCR Enzyme Mix	2.0
Template RNA	4.0
RNase-free water	18.0
Total	50

4 Conclusion

Among RNA viruses, Rubella has the highest GC content (approx 70%) and therefore the formation of a secondary structure during amplification should be considered during the setting of RT-PCR reaction. For the 480 bp region GC% accounts for 67.8% and for the 633 bp region it is 70%. Using the Superscript III one step RT-PCR Kit in the absence of betaine (1 M) and further analysis with agarose gel analysis of the PCR product showed no difference in the molecular weight of both products (fragment-1—480 bp region and fragment-2—633 bp region). Analysis of the sequence of the fragment-2 PCR product showed that a region of about 150 bp deletion occurred towards the 5′ end of the sequence. The length of this deletion was variable for different samples. The enzyme reverse transcriptase is used for in vitro reverse transcription of RNA to its complementary DNA and is the most important step in genotyping and gene expression studies. Thus in our experience and under the specific conditions mentioned in the method section during genotyping of rubella with reverse transcriptase, caution should be exercised to interpret the results. Due to the high GC content in the genome and the propensity to form secondary structures by RV genome deletion of specific region could be observed while using the Superscript III enzyme in the absence of betaine. The protocol discussed in this chapter is also recommended by WHO and for further details, the following reference is recommended [15].

5 Notes

1. Inclusion of betaine 1 M (Final concentration) is a necessary condition during genotyping using the superscript III RT-PCR kit. The betaine used should be procured separately as it is not a part of the kit. We have observed deletion of variable length in fragment 2 during genotyping using superscript III in the absence of betaine (1 M). In case of Qiagen Kit, the 5x Q-Solution provided with the kit should be used.

References

1. Matonaviridae ~ ViralZone page (https://viralzone.expasy.org/8916?outline=all_by_species)
2. Chen R, Mukhopadhyay S, Merits A, Bolling B, Nasar F, Coffey LL, Powers A, Weaver SC (2018) ICTV virus taxonomy profile: Togaviridae. J Gen Virol 99:761–762
3. Parkman PD. Togaviruses: Rubella Virus. In: Baron S, editor. Medical Microbiology. 4th edition. Galveston (TX): University of Texas Medical Branch at Galveston; 1996. Chapter 55. Available from: https://www.ncbi.nlm.nih.gov/books/NBK8200/
4. Bennett AJ, Paskey AC, Ebinger A, Pfaff F, Priemer G, Höper D, Breithaupt A, Heuser E, Ulrich RG, Kuhn JH, Bishop-Lilly KA, Beer M, Goldberg TL (2020) Relatives of rubella virus in diverse mammals. Nature 586:424–428
5. Witte JJ (1969) Epidemiology of rubella. Arch Pediatr Adolesc Med 118:107
6. Achievements in Public Health: Elimination of Rubella and Congenital Rubella Syndrome United States, 1969–2004
7. Bartlett JMS, Stirling D (2003) A short history of the polymerase chain reaction. Methods Mol Biol 226:3–6. https://doi.org/10.1385/1-59259-384-4:3.
8. Zhou Y, Ushijima H, Frey TK (2007) Genomic analysis of diverse rubella virus genotypes. J Gen Virol 88:932–941
9. Frey TK (1994) Molecular biology of rubella virus. Adv Virus Res 44:69–160
10. WHO (2013) Rubella virus nomenclature update: 2013. Wkly Epidemiol Rec Relev épidémiologique Hebd 88:337–343
11. Zheng D-P, Frey TK, Icenogle J, Katow S, Abernathy ES, Song K-J, Xu W-B, Yarulin V, Desjatskova RG, Aboudy Y, Enders G, Croxson M (2003) Global distribution of rubella virus genotypes. Emerg Infect Dis 9:1523–1530
12. Abernathy E, Cabezas C, Sun H, Zheng Q, Chen M-H, Castillo-Solorzano C, Ortiz AC, Osores F, Oliveira L, Whittembury A, Andrus JK, Helfand RF, Icenogle J (2009) Confirmation of rubella within 4 days of rash onset: comparison of rubella virus RNA detection in Oral fluid with immunoglobulin M detection in serum or Oral fluid. J Clin Microbiol 47:182–188
13. Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549
14. Staden R (1996) The staden sequence analysis package. Mol Biotechnol 5:233–241
15. WHO. Rubella genotyping RT-PCR 2-fragment, Invitrogen (CDC). https://www.who.int/immunization/monitoring_surveillance/burden/laboratory/AnnexC3_Rubella_Genotyping_RT-PCR_Superscript.pdf?ua=1