Laboratory Exercises

An Exercise in Molecular Epidemiology* &
HUMAN RHINOVIRUS PREVALENCE AND GENETICS

Received for publication, March 16, 2011 and in revised form, March 16, 2011

Catherine J. Albright and David J. Hall
From the Chemistry Department, Lawrence University, Appleton, WI 54911

Human rhinovirus (HRV) is one of the most common human respiratory pathogens and is responsible for the majority of upper respiratory illnesses. Recently, a phylogeny was constructed from all known American Type Culture Collection (ATCC) HRV sequences. From this study, three HRV classifications (HRVA, HRVB, and HRVC) were determined and techniques for classifying new isolates of HRV were reported. The genetic change of this virus in specific populations over time is of great interest to understand the evolution and epidemiology of viruses. To facilitate the collections of HRV sequences over a number of years, a virology experiment was designed in which students test nasal lavage samples to look for HRV infection. Students will learn a variety of techniques including RNA isolation, cDNA synthesis, qPCR, and agarose gel electrophoresis as well as bioinformatic skills though examination of sequences from the HRV-field isolates. Furthermore, students can look at symptom data from subjects to investigate correlations between symptom severity and factors such as stress and sleep patterns. Such information can be used to examine hypotheses regarding HRV mutation, symptom severity and epidemiology.

Keywords: qPCR, virus, genomics, RNA.

INTRODUCTION

Virology is an interdisciplinary field that incorporates aspects of biochemistry, cell biology, molecular biology, and immunology. Thus, the study of virology is intriguing to students not only because of its depth but also because of viruses that have a very visible effect on a university population when cold and flu season arrives. Viruses are parasitic nucleic acids, either DNA or RNA, and function through the “hijacking” of the host cell’s own machinery. A productive viral infection must spread virus to other hosts to continue the viral life cycle, and the study of viral spread in a population is a principle concern of viral epidemiology. Viral epidemiology not only examines how a virus functions in the host but also opens a window into human habits and interactions. Some viruses are transmitted through aerosols, others via an oral/fecal route, and still others through the transmission of bodily fluids [1]. Thus, many viruses thrive by taking advantage of host habits. By understanding how viruses replicate and spread, students can gain insight into the functions of viruses, cells, and evolution. Many viruses quickly adapt to environmental changes through error-prone replication, resulting in mutations and the generation of viral quasi species in which many genotypes coexist simultaneously during replication in the host cells. Subsequent infection in a new host results in replication by viruses fit for a particular environment [1, 2].

Human rhinovirus (HRV), a picornavirus, is a member of the largest pathogenic virus family that includes such members as poliovirus, coxsackie virus, hepatitis A, and foot and mouth disease virus. Picornaviruses are positive (mRNA-like), single-stranded RNA viruses, encased in a 30-nanometer icosahedral protein capsid. Their genomes contain roughly 8,000 nucleotide bases that code for 12 proteins [1, 3]. HRV has a well-characterized genome with three main classifications currently identified: HRVA, HRVB, and HRVC. During replication, millions of viral particles are made; however, the nucleic acids are not exactly identical. The HRV polymerase is error prone, and so several mutations occur through each round of replication. Thus, each viral nucleic acid can be slightly different. These variants are referred to as viral quasi species and enable HRV to adapt to the host [1, 2].

To date, there are 99 HRV serotypes recognized by the American Type Tissue Culture Collection (ATCC), although there are likely many more in the population considering the viral quasi species phenomenon. HRV infections occur worldwide at all times of the year, with...
strong outbreaks in the fall and late spring. However, HRV is only one of many respiratory viruses—others, including adenovirus and coronavirus, produce similar symptoms [1, 3]. HRV infections are usually limited to the upper respiratory tract, though they can also result in acute otitis (inflammation of the middle ear), sinusitis (inflammation of the paranasal sinuses), and lower respiratory tract infections [4–6].

The extensive data available on HRV genomes enables many bioinformatics opportunities for students, including alignment of genome sequences to look for mutations at the RNA level and differences among protein sequences. Students can examine differences in the HRV serotypes over several years of data regarding a university population to identify HRV mutations that have occurred and their severity in causing symptoms in the host. This investigation is not only a way for students to study virus evolution but serves as an introduction in viral epidemiology.

In this inquiry-based laboratory project, students use a variety of techniques, including RNA isolation, cDNA synthesis, qPCR, and agarose gel electrophoresis to look for the presence of HRV in nasal lavage samples from human subjects. The samples that are shown to be positive for HRV can then be sequenced to determine the strain of the isolate. By analyzing surveys in which subjects indicate severity of their symptoms, stress factors, and average hours of rest per night, students can identify possible contributors to HRV infection. These factors have been seen to make subjects more susceptible to the common cold and other viruses that cause upper respiratory infections [7–10].

**LABORATORY EXERCISE**

Students who participate in this exercise should have basic biochemistry and molecular biology laboratory skills. They should be familiar with basic laboratory equipment (pipettors and centrifuges) and agarose gel electrophoresis if there is no qPCR thermocycler available to the lab. The laboratory exercise has convenient stopping points and can be completed in two laboratory periods if done efficiently. This laboratory works well with teams of two students who are allowed to divide the work.

**Session 1**

• Students collect nasal lavage samples from themselves and classmates and survey their symptoms—this can be done before lab if student are equipped with needle-less syringes and sterile tubes and instructed in collection technique.

• Students isolate RNA from samples and synthesize cDNA (process can be stopped after isolation, but it is best to proceed to cDNA step if possible).

**Session 2**

• Students setup and perform both qPCR and PCR—while these are running, students can analyze data from symptom surveys.

• (Students can run an agarose gel of the PCR product and image it.)

**Session 3**

• (Students can run an agarose gel of the PCR product and image it.)

• Students prepare samples for sequencing.

**Session 4—when sequence data is returned**

• Students analyze sequences and draw conclusions regarding differences in serotypes found and how they differ from previous years.

**Ethics and Safety**

All aspects of the study involving human subjects must be approved by the school’s institutional review board. Subjects must be informed of the experiment and that use of the data and must sign an approved consent form before a sample can be obtained from them. Students in the laboratory must be informed of the importance of anonymity of the samples they are testing, which should be assigned initials of the subject or, better yet, numbers.

Students should wear gloves whenever they are handling samples or cDNA. Special care should be taken when using TRIzol reagent—TRIzol should be used in a chemical fume hood to avoid inhalation of vapors. Students should wear gloves and protective eyewear, avoiding any chemical contact with the skin or clothing.

**MATERIALS AND METHODS**

**Nasal Lavage—Collect Mucosal Samples from a Variety of Sick and Healthy Students**

Students first collected nasal lavage samples from themselves and their classmates. Subjects were informed of the experiment and signed a consent form. Students used a 1-mL syringe (Becton Dickinson, Franklin Lakes, NJ) to inject 1 mL of water into the nostril of the subject by placing the tip of the syringe just inside the nostril touching the inside of the nose and injecting the water. The water was then allowed to fall back out of the nose and collected in a sterile 15-mL centrifuge tube (Cellstar–Greiner Bio-one, Monroe NC) held under the nostril. Students also asked their subjects to fill out a short survey (attached along with consent form) regarding their health, sleep habits, and stress levels. Subjects rated symptoms associated with the common cold on a scale of 0 to 4, with 0 being non-existent and 4 being severe (survey included, see Supporting Information 3).

**RNA Isolation and cDNA Synthesis**

Filtered pipet tips and gloves were used at all times to prevent RNA degradation of the sample. From each sample, 350 μL were moved to a sterile 1.8-mL microfuge tube (Eppendorf, Hamburg, Germany) with 25-μL 5× nucleic acid extraction buffer (Final concentrations of components: pH7.5 Tris 500 mM, NaCl 1500 mM, and EDTA 5 mM) (Supporting Information 2) and 10 μL of carrier mix for RNA isolation consisting of 0.5 μL of human DNA (100 ng/μL) (Sigma-Aldrich), 4 μL glycogen (5 μg/μL) (Applied Biosystems, Carlsbad, CA), 1 μL of glycoblue (15 μg/μL; Applied Biosystems, Carlsbad, CA) and 4.5 μL of phosphate buffered saline. RNA was isolated from the sample by adding 750 μL of TRIzol LS Reagent (Invitrogen, Carlsbad, CA), and the tubes were vortexed (Scientific Industries Inc, Bohemia, NY) vigorously for 1 minute before 200 μL chloroform was added. The tubes were again vortexed for 1 minute and
then centrifuged (Eppendorf Centrifuge 5415C, Hamburg, Germany) at 13,000 × g for 10 minutes. The aqueous layer was removed to a new Eppendorf tube containing 600 μL of RNase-free isopropyl alcohol. The tubes were then inverted several times, allowed to stand for at least 5 minutes at room temperature and then centrifuged at 13,000 × g for 10 minutes to pellet the RNA. The supernatant was then removed and 700 μL of 75% ethanol, at room temperature, were added to the pellet. The tubes were then centrifuged for another 5 minutes at 13,000 × g, and the supernatant was carefully removed so as not to disturb the pellet of RNA. Each RNA pellet was then resuspended in 22 μL RNase-free water and vortexed vigorously to dissolve the RNA.

Reverse transcription was conducted by first placing a mixture of the following reagents in each 0.5-mL Eppendorf tube: 8 μL 5× AMV-RT buffer, 8 μL 5 mM dNTP (Promega, Madison, WI), 1.0 μL random hexamers (IDT, Coralville, IA), 0.5 μL RNAse inhibitor (Promega, Madison, WI), 0.6 μL AMV (Promega, Madison, WI) reverse transcriptase, and 5.9 μL RNase-free water. Next, 16 μL of viral RNA were added to each tube. An upscaled master mix was prepared and apportioned for multiple samples. Reverse transcription was performed in a thermocycler (MJ Research, Waltham, MA) with the following program: 25 C for 5 minutes, 42 C for 10 minutes, 50 C for 20 minutes, and 85 C for 5 minutes.

**qPCR—One Method of Looking for Viral Presence in Collected Samples**

qPCR was performed on the cDNA from samples (Applied Biosystems 7500 Real Time PCR System, Carlsbad, CA) using FastStart Universal SYBR Green Master (ROX) mix (Roche, Biosystems 7500 Real Time PCR System, Carlsbad, CA) using the noncoding region of HRV [5]. The primer stocks were resuspended in water at a concentration of 200 μM each. A 150-μL mix of the forward primer mix was made by adding 18.75 μL of each P1-1 primer to 93.75 μL water for a final concentration of 25 μM for each primer (Supporting Information 2). The P3-1 primer was diluted 1:8 to a final concentration of 25 μM. A 20-μL reaction was performed using 1 μL of cDNA, 10 μL of the SYBR Green Master mix, 1 μL of forward primer mix (P1-1, B1: CAAGCACTTCTGTTCCCACTCCC B2: CAAGCACCCTCTGTTCCCACTCCC B3: CAAGCACTTCTGTTCCCACTCCC B4: CAAGCACCCTCTGTTCCCACTCCC B5: CAAGCACCCTCTGTTCCCACTCCC) and 1 μL of the reverse primer (P3-1: ACGGACCCCCAAGTAG; see Supporting Information 2). An upscaled master mix was prepared and apportioned for multiple samples.

**PCR and Agarose Gel Electrophoresis—Alternative Way of Determining Presence of Virus**

Two rounds of PCR amplification were performed using (P1-1, B1: CAAGCACTTCTGTTCCCACTCCC B2: CAAGCACCCTCTGTTCCCACTCCC forward primer mix and P1-3 reverse primer (ACGGACCCCCAAGTAG). The primer stocks were resuspended in water at a concentration of 200 μM each. A 150-μL mix of the forward primer mix was made by adding 18.75 μL of each P1-1 primer to 93.75 μL water for a final concentration of 25 μM for each primer. The P3-1 primer was diluted 1:8 to a final concentration of 25 μM. The PCR reaction contained 22.5 μL of the PCR Platinum SuperMix High Fidelity (Invitrogen, Carlsbad, CA) mix, 1 μL of P3-1 primer (ACGGACCCCCAAGTAG), 0.5 μL of the P1 mix, and 2.5 μL of cDNA for a total of 26.5 μL. The reaction was run with a touchdown PCR according to the reaction conditions in Supporting Information 2 [5].

This first reaction was followed by a second PCR: The PCR contained 45 μL of the Platinum mix, 1 μL of P3-1 primer (25 μM), 1 μL of the P1-1 (25 μM), and 3 μL of the products from the previous PCR. The products were analyzed in the following way: A 1.5% agarose gel was used to separate the products of the PCR (SeaKem LE, Erie, PA; Phusion, NEB, Ipswich, MA) Hi-Lo® DNA maker loading dye (Promega, Madison, WI). Gels were made with 1× Tris Borate EDTA (TBE) buffer and run at 80V (VWR Power Source 250 V, West Chester, PA). Nucleic acids were imaged using ethidium bromide (EtBr) and a Fotodyne system (Hartland, WI).

**Sequencing**

HRV PCR products were sequenced to determine the genetic makeup of each field isolate. Samples were prepared by adding 4 μL of Exo-SAP-IT (USBiochemicals, Cleveland, Ohio) into 10 μL of PCR product. This mixture was incubated at 37°C for 25 minutes and then 90°C for 15 minutes. The resulting product was stored at -20°C until sequencing could be performed. Sequencing was performed by DeWalch Technologies, (Houston, TX) using primer P3-1 (ACGGACCCCCAAGTAG; see Supporting Information 2).

**Multiple Alignment and Phylogeny**

Sequences from the samples were entered into BLAST [11]. The resulting BLAST information established to which HRV sequence from the ATCC the sample was most closely related. A multiple sequence alignment and cladogram (using group tree) were both generated using web-based ClustalW2 following the instructions provided [12].

**Analysis of Stress, Sleep, and Symptom Levels of Subjects**

Questionnaires for determining a subject’s stress sleep and symptom severity were adapted from previously published studies [7, 8, 13, 14] (Supporting Information 3). Briefly, stress was assessed through a subject’s rating his or her average stress level on a 1–5 scale with 5 being the most severe. The quality and duration of sleep was assessed based on the reported number of hours (1–10). The severity of symptoms was determined by subjects rating each of their cold symptoms on a scale of 0–4 (4 most severe) and adding the resultant scores together to give a final symptom score. The symptom score was then adjusted to a 1–10 scale. A two-tailed t test was then used to determine the significance (GraphPad Software, Inc http://www.graphpad.com/quickcalcs/index.cfm).

**RESULTS AND DISCUSSION**

**HRV Detection**

The primer set used to detect HRV in the samples was adapted from Lee et al. [5] and detects all 99 possible ATCC serotypes through the amplification of ~400-bp portion of the 5′ noncoding region of HRV. The detection of HRV was conducted with both conventional PCR and qPCR using the same primer set. The qPCR method is preferable due to increased sensitivity. Great care needs to be taken when preparing the viral RNA from samples to prevent degradation and loss of the RNA pellet during the precipitation procedure.

Of the 28 samples obtained from human volunteers, the qPCR indicated 17 were positive for HRV (Fig. 1). The 17 positive samples showed clear differences in the amount of viral nucleic acid present as indicated by the differences in cycle number. These differences can be manifest for several reasons including stage of the infection and the efficiency of HRV nucleic acid isolation. One
can quantify the exact amount of virus using laboratory-grown strains, but this approach is beyond the scope of this article and the availability of equipment in many undergraduate laboratories.

The standard PCR and subsequent agarose gels demonstrated this technique’s weaker sensitivity compared to the qPCR (Figs. 2a and 2b). Of the 17 positives on qPCR, six were positive with standard PCR. Samples with low amounts of viral nucleic acid or samples contaminated with RNases likely only appear as positive on the qPCR. No samples appeared as positive in conventional PCR after a negative result from qPCR.

Sequencing

There are 99 HRV serotypes recognized by the ATCC and extensive data available on HRV genomes enables many bioinformatics opportunities for students, including alignment of genome sequences to look for changes at the RNA level. Standard PCR was used to generate enough genetic material for sequencing, and students retrieved sequences from the service provider website and submitted them as BLAST queries to identify to which ATCC strain their HRV samples were most similar [11]. The analysis revealed that two of the six samples were variants of HRV1 with 97% of the sequence identical with the ATCC reference strain. The other samples were identified as being similar to ATCC reference strains HRV2, HRV16, and HRV85. The sequences were aligned with each other and ATCC reference strains using ClustalW2 and a group tree cladogram was generated (Fig. 3) [12]. From a research perspective, the sequence data is most interesting because, as samples are sequenced yearly, it will be possible to create a genetic database of HRV strains found in a college-aged population and to do further analysis on many levels. Indeed, a discussion was initiated with the students on how the nucleotide variations between their samples and the ATCC strains may influence the ability of the virus to elude the immune system, replicate or spread between hosts.

Additional avenues of investigation can include the analysis of predicted RNA secondary and tertiary structure [6, 15] and the sequencing of the nucleic acid responsible for coding of the viral capsid or the entire genome. With this information and a free program called VIGOR, students can identify all 12 viral proteins and perform bioinformatics analysis in comparison to ATCC reference strains [16]. In a sense, this laboratory is similar to the program sponsored by Howard Hughes Medical Institute called “Phage-Hunters” [17].

Fig. 1. Quantitative PCR of nasal lavage samples. qPCR spectra indicating HRV presence in samples. Twenty-eight total samples were run—14 of subjects who reported being sick with common cold symptoms and 14 who reported no symptoms. The x-axis indicates the number of qPCR cycles, and y-axis indicates the relative fluorescence units. Dashed lines are those subject who were HRV positive but did not report being sick. The qPCR was as described in Methods and was performed by Lawrence University students during an undergraduate general biology class.

Fig. 2. Standard PCR of nasal lavage samples. PCR was conducted as described and products were subjected to gel electrophoresis and stained with SYBR green. (a) Agarose gel indicating no successful HRV standard PCR in lanes 1, 5, 9, 10. Successful standard PCR amplification for sequencing of qPCR samples in lanes 2, 3, 4, and 6. Lanes 7 and 8 correspond to HRV1 and HRV 2 positive controls, respectively. (b) Agarose gel indicating no successful HRV standard PCR in lanes 1, 2, 3, 6, 7, 9, and 10. Successful standard PCR amplification for sequencing of qPCR samples in lanes 4 and 5. Lane 8 corresponds to the HRV16 positive control. Standard PCR and subsequent agarose electrophoresis was done as described in Methods and was performed by Lawrence University students during an undergraduate general biology class.
Correlation Between Symptoms, Sleep, Stress, and HRV Infection

Studies have shown that the severity of experimental viral infections depends not only on the viral strain but also on the sleep and stress levels of the individuals involved [7–10]. In this study, 21 of the 28 students volunteering samples chose to participate in the symptoms, stress, and sleep surveys as well (Figs. 4a–4c). Only 10 of the 17 HRV positive subjects chose to participate in the surveys while all 11 of the subjects that tested negative for HRV participated in the surveys. Of those subjects negative for HRV, five were found to have reported being sick. In this instance, other viruses are probably responsible for the observed symptoms. This laboratory can easily be extended to include PCR primers to identify other viruses including influenza, corona virus, and adenovirus among other possibilities [18].

The average symptom, sleep, and stress scores were compared between HRV positive and HRV negative subjects and between reported sick and healthy subjects (Figs. 4a–4c). In our data, there are two significant trends as determined by an unpaired t test. First, symptom scores correlated positively with being HRV positive ($p = 0.15$, not significant) or sick ($p = 0.0001$, significant; Fig. 4a). Second, the sleep amount correlated positively with being HRV negative ($p = 0.1$, not significant) or healthy ($p = 0.0077$, significant; Fig. 4b). However, there was no apparent relationship between self-reported stress level and having HRV or being sick (Fig. 4c). A larger $N$ value can be achieved if this laboratory is taught over multiple years, and data is analyzed in aggregate.

In addition to looking for correlations between symptoms, stress levels, amount of sleep and infection, students can collect information regarding subject residences and activity participation. Students can group subjects by dormitory and when the results of the sequencing return, see whether or not different residences on campus harbor different HRV strains. However, great care must be taken to ensure student privacy in the data collection and encoding process.

CONCLUSIONS

The value in doing this laboratory is not only pedagogical; in teaching it yearly and sequencing to track the infectious cycle of HRV, it will be possible over the years to build a record of what serotypes of HRV are affecting a population, with particular symptoms, at a given location. This information can help in answering questions such as: What types of mutation occur? What types of HRV are prevalent? What percentage of the respiratory infections is HRV? Who contracts HRV? How does this correlate to sleep habits and stress levels? By creating such a dataset, insight can be gained regarding HRV infection and mutation in a college population.
Potential Student Pitfalls

RNA degradation is the single most troublesome area. Although positive and negative samples can easily be identified via qPCR, getting enough PCR products for the detection by gel electrophoresis can be hampered by RNA degradation. RNase-free reagents should be used at all times, and gloves are a must. Furthermore, delays in the procedure for isolating the RNA can result in degradation. Often, running a third PCR can be helpful in these situations. This extra step can be done by using 45 μL of the Platinum mix, 1 μL of P3-1 primer (25 μM), 1 μL of the P1-1 (25 μM), and 5 μL of the products from the previous PCR.

Acknowledgment—The authors thank Sally Berben, Ronald F. Peck, Michael Schreiber, and Wai-Ming Lee for technical assistance and critical reading of the manuscript.

REFERENCES

[1] S. J. Flint, L. W. Enquist, V. R. Racaniello (2009) Principles of Virol-ogy, 3rd ed., ASM Press, New York.
[2] A. L. Kistler, D. R. Webster, S. Rouskin, V. Magrini, J. J. Credle, D. P. Schnurr, H. A. Boushey, E. R. Mardis, H. Li, J. L. DeRisi (2007) Virol. J. 4, 40.
[3] T. J. Tuthill, E. Groppelli, J. M. Hogle, D. J. Rowlands (2010) Curr. Top. Microbiol. Immunol. 343, 43–89.
[4] D. Kiang, S. Yagi, K. A. Kandtardjieff, E. J. Kim, J. K. Louie, D. P. Schnurr (2007) J. Clin. Virol. 38, 227–237.
[5] W. M. Lee, C. Kiesner, T. Pappas, I. Lee, K. Grindle, T. Jartti, B. Jakiela, R. F. Lemanske, Jr, P. A. Shult, J. E. Gern (2007) PLoS One 2, e966.
[6] A. C. Palmenberg, D. Spiro, R. Kuzmickas, S. Wang, A. Djikeng, J. A. Rathe, Fraser-C. M. Liggett, S. B. Liggett (2009) Science 324, 55–59.
[7] S. Cohen, W. J. Doyle, C. M. Alper, D. Janicki-Deverts, R. B. Turner (2009) Arch. Intern. Med. 169, 62–67.
[8] S. Cohen, W. J. Doyle, R. Turner, C. M. Alper, D. P. Skoner (2003) Psychol. Sci. 14, 389–395.
[9] S. Cohen, D. A. Tyrrell, A. P. Smith (1993) J. Pers. Soc. Psychol. 64, 131–140.
[10] S. Cohen, D. A. Tyrrell, A. P. Smith (1991) N. Engl. J. Med. 325, 606–612.
[11] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman (1990) J. Mol. Biol. 215, 403–410.
[12] M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGet- tigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, D. G. Higgins (2007) Bioinformatics 23, 2947–2948.
[13] S. Cohen, C. M. Alper, W. J. Doyle, N. Adler, J. J. Treanor, R. B. Turner (2008) Health Psychol. 27, 268–274.
[14] S. Cohen, C. M. Alper, W. J. Doyle, J. J. Treanor, R. B. Turner (2006) Psychosom. Med. 68, 809–815.
[15] A. C. Palmenberg, J. A. Rathe, S. B. Liggett (2010) J. Allergy Clin. Immunol. 125, 1190–1199.
[16] S. Wang, J. P. Sundaram, D. Spiro (2010) BMC Bioinformatics 11, 451.
[17] S. M. Caruso, J. Sandoz, J. Kelsey (2009) CBE Life. Sci. Educ. 8, 278–282.
[18] W. M. Lee, K. Grindle, T. Pappas, D. J. Marshall, M. J. Moser, E. L. Beaty, P. A. Shult, J. R. Prudent, J. E. Gern (2007) J. Clin. Microbiol. 45, 2626–2634.