Virus discovery and human parechoviruses

de Vries, M.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Summary

In 5-30% of children that are hospitalized with respiratory illness a viral pathogen is suspected but the agent remains unidentified. It is possible that these children are infected with a yet unknown virus. To detect a virus with molecular techniques such as PCR and sequencing, one could use a sequence dependent or a sequence independent method. In case a sequence dependent method is used an indication of a virus-family is necessary. For a sequence independent method (like Virus Discovery cDNA-AFLP (VIDISCA) and Random priming) no sequence knowledge is necessary, but these methods are hampered by background DNA/RNA (e.g. ribosomal RNA). Therefore it is almost not possible to use these methods for patient material because it contains too high concentrations of background competitor nucleic acids, mainly ribosomal RNA (rRNA). The sequence independent methods can easily be used in virus culture supernatant but unfortunately not all viruses can be cultured. To identify these viruses the sequence independent methods should be optimized such that identification in clinical material is possible.

In chapter 2 the optimization of VIDISCA is described to allow viral identification directly in clinical specimen. Using 96 hexamer primers that anneal to all known viruses but not to rRNA, VIDISCA became 100x more sensitive. By changing the restriction enzymes (from HinP1 + Mse1 to only Mse1) we increased the possible viral fragments that can be amplified and simultaneously reduced the number or rRNA fragments. The remaining rRNA fragments are further inhibited by rRNA-blocking oligonucleotides that are added during reverse transcription. These oligonucleotides cannot be extended at the 3’ end and therefore block reverse transcription of the fragments to which they anneal to. Besides the reduction of background rRNA amplification the VIDISCA sensitivity was increased further by using high throughput sequencing, allowing a minority population to be identified if a sufficient number of sequences are generated for the VIDISCA amplified fragments. With VIDISCA-454 we were able to identify the virus in 11 out of 18 clinical respiratory samples. A lower limit of detection could be determined which is approximately 10,000 viral genome copies per ml, which is extremely sensitive for a sequence unspecific amplification technique.

The high sensitivity which was measured in chapter 2 was determined in respiratory material. In chapter 3 the sensitivity in blood and stool samples was examined. In blood-serum the sensitivity equals the sensitivity in respiratory material, thus also 1,000 viral genome copies as input (100 µl). In stool suspensions the background amplification caused some decrease in sensitivity, and
approximately 10,000 viral copies are needed for detection. For both types of material these concentrations are easily met during an acute virus infection.

In chapter 4 a search for a virus involved in canine hepatitis was performed. VIDISCA-454 easily showed a canine adenovirus to be present in a virus culture supernatant of a dog with hepatitis. In total we could identify 7738 viral reads out of a total of 9190 (84%) which shows the strength of the VIDISCA-454 protocol.

In chapter 5 a search for unknown viruses in unidentifiable virus cultures was performed, which resulted in the identification of human parechovirus (HPeV) 5 and 6. HPeV-5 was not identified yet in the Netherlands and HPeV-6 was still unknown at the time of identification, but a group from Japan described the virus shortly after we obtained the first viral sequences.

Multiple novel HPeV variants have been discovered in the last few years. To determine whether these variants might represent new introductions in the human host we examined HPeV evolution, and calculated the time to the most recent common ancestor (TMRCA) in chapter 6. For HPeV the nucleotide substitution ratio of the P1 region (N=29; median 2.21 x 10^{-3} s/s/y) and VP1 (N=199; median: 2.79 x 10^{-3} s/s/y) was determined. The TMRCA was estimated to have existed around year 1600 by using the sequences of VP1 (N=199), thus it seems likely that the parechoviruses were first introduced in human about 400 years ago.

In chapter 7 the genetic relationship between the various HPeV strains was determined with 18 full-length genomes and those deposited in the GenBank database. Our analyses suggest that type HPeV-1 can be divided into HPeV-1A and HPeV-1B. The sequence variation in the VP1 gene (which is used for typing HPeV) between the two HPeV-1 clades is of a magnitude that is between that of inter- and intra-species comparisons. Furthermore we identified possible recombination points in the HPeV genomes which are mostly located at the 5’ and 3’ end of the P1 region.