Reliability of PCR to detect genetic sequences from HIV

Juan Manuel Morillo Velázquez

Abstract

Polymerase chain reaction (PCR) is a lab technique used for detection of genetic sequences derived from microbial or human genes, so it can be used in microbiological and genetic diagnosis. Its applications in HIV/AIDS has been increasing in the last 15 years, in viral load testing to measure the number of copies of HIV RNA in human plasma, and for detection of HIV DNA (proviral DNA) within human cells. However, there are several studies reporting the low specificity of the method. One of the reasons could be the high genetic diversity of HIV, but other reason could be the presence of endogenous retroviral sequences in human samples. An approach to test the reliability those sequences usually employed in commercial tests, and to study the hypothesis that genetic sequences considered as HIV could be endogenous is presented.
What is PCR?

Polymerase chain reaction (PCR) is a molecular technique used generally to obtain multiple copies from a specific fragment of microbial or human DNA, in order to detect a specific microorganism (in the case of microbiological diagnosis), or a variation in the consensus sequence of a human gene (in the case of genetic conditions, such as cystic fibrosis).

This method is performed totally in vitro. For diagnosis, it is a prerequisite to know previously a part or the whole of the genome sequence of the microorganism to detect, or the human gene to test. Then, it is necessary to select a specific region of that sequence that doesn’t cross-react with other sequences from other microorganism or genes. Currently, there are several bioinformatics tools to perform this task (an example is BLAST) from several database of genetic sequences available (for instance, GenBank).

Once the researcher chooses the target region, he/she has to design two short fragments of DNA that have complementary nitrogen bases of both extremes of the chosen target sequence. These short molecules are called primers, and they are the specific component in this chemical reaction. The rest of chemicals are similar, in general terms, independently of the sequence and its source organism (microorganism or human). These chemical are the following:

- **Template DNA**: in the case of medical diagnosis, it is the whole DNA in a sample derived from biological fluids or tissues, and its origin can be microbial or human.

- **Nucleotides**: these are each basic piece in a DNA molecule. There are four types, according the nitrogen base: ddATP (with adenine), ddGTP (guanine), ddTTP (thymine), ddCTP (cytosine).

- **DNA-polymerase**: it is an enzyme that binds the DNA template when a primer binds to this one (because of base complementarity), and then starts joining new nucleotides to the primer, according the sequence of the DNA template. It is the “motor” of PCR.

Usually, all the chemicals are mixed in water-based solution with adjusted pH, and the procedure depends on several temperature changes. When the mix is at 95°C, template double-strand DNA can be denaturated, and each strand is separated from the other. Then, temperature decreases to 55-60°C, when primers are annealed to the template DNA, whenever they recognize the specific sequence because it exists within the template. The last step is to increase the temperature to 72°C, when polymerase can function more efficiently and produce new copies. These three steps are called a cycle, and the procedure usually consists of 30-40 cycles. This can be performed automatically in a machine called thermocycler.
The final product of the procedure must be detected in order to assess the reliability and precision of the technique. The simplest method to visualize PCR products is using a procedure called electrophoresis. It consists in putting the post-amplification mix into a solid support, such as an agarose gel (soft but firm), and separate each amplified molecule according its size when an electric current is applied to this gel. The internal structure of the gel is a mesh, and DNA molecules runs to the opposite end because of the current, but larger molecules are slower and smaller molecules are faster, so they can be separated and recognized according their predicted size. Then, gel must be stained in order to visualize the pool of molecules with similar sizes. This usually is performed by using ethidium bromide. Hence, we could check if the selected sequence (with a known size) exists in the analyzed sample, or if it was no present or in small amounts, insufficient to visualize.

Theoretically, we only should see one DNA band in the gel, because PCR should have been specific. But frequently it is possible to find multiple bands in case of:

- Too low temperature for primers annealing: when this occurs, primers can bind to other regions of template DNA with a not totally complementary sequence, so mismatching has occurred. Hence, PCR is very sensitive to small changes in temperature and chemical concentrations, and even thermocyclers from different brands adjusted with the same temperature profile yield different results.

- Unknown sequences in the sample: although a great amount of genetic sequences are currently known and registered in open access database, there are huge amounts of unknown microorganisms or variants from them that makes very difficult to have a high level of reliability that those sequences detected in a PCR experiment belong really to the searched species. Moreover, in the sample there are plenty of sequences derived from RNA (ribonucleic acid) formed during cellular function that could crossreact with primers and yield unspecific bands.

In the practice, PCR is more useful and reliable when human sequences are used for genetic diagnosis than when it is used for microbiological testing. In this case, simple variations in the sequence of a microorganism could lead to a non detection of a sequence (false negative result) or the binding to sequences of other close related microorganisms (false positive result).

In the last years, several approaches to increase the reliability of detection of specific sequences have been reported and have improved the applicability of PCR in clinical diagnosis. Nevertheless, standardization of commercial PCR kits for microbiological testing is difficult, and caution should be taken into account in order to not relying a clinical decision only on PCR results.
PCR and HIV

Since antibody testing for HIV can give not totally reliable results, and are not early detected (according to the official theory), PCR was introduced in HIV research since early 1990s. In HIV research, PCR can be used to detect RNA (the original molecule within HIV particles) or DNA (the molecule derived from reverse transcription of RNA that is hosted in human cells). In mid 1990s, viral load testing was introduced for the follow up of HIV/AIDS patients, and this meant that the number of RNA HIV copies was measured in the plasma, and this result correlated with prognosis, according to the official theory. Several methods have been developed for viral load measuring, but the main has been RT-PCR (reverse transcription-PCR). Reverse transcription is a natural process that implies the synthesis of DNA from RNA, and it is performed by an enzyme called reverse transcriptase. Forty years ago, researchers thought that this enzyme was only present in retrovirus, but currently it is accepted that mostly organisms have reverse transcriptases, whose function is not totally elucidated. When DNA has been obtained from RNA, a modified PCR is carried out, in order to quantify the final number of copies. Usually, this has been obtained by using a labeled DNA fragment complementary of the central region to be amplified. This labeled DNA could be employed to produce a color reaction, in such manner that the intensity of the color was proportional to the amount of final product. Although this procedure has been used for more than 10 years, it has low precision, because there are multiple factors influencing the yield of the PCR reaction.

Currently, the procedure that has an increasing acceptation is real time PCR. This technique has higher control of factors influencing PCR, and products can be detected during the process, not only at the end. It is more reproducible, but some discrepancies have been described depending the HIV sequence selected for diagnosis, and commercial kit used. Thus, the main manufacturers, Roche and Abbott, are competing in order to impose their respective kit in the market.

In these years, the high genetic diversity of HIV has been considered a handicap to standardize PCR testing. In 1998, a paper from one of the main Spanish hospitals attending HIV/AIDS subjects found near 20% of false positive results with PCR testing when they used a Roche kit with primers designed to detect a great amount of HIV subtypes. This diversity has also been attributed the discrepancy of results derived from PCR with HIV DNA sequences.

HIV DNA is detected to study latent infection of human cells. Theoretically, although viral load (RNA) was not detected in a HIV+ subject, HIV DNA should be detected so HIV cannot be totally eliminated according to the official theory. In the practice, several papers published in the last 15 years report discrepant results between serological techniques and PCR DNA, with false negative and false positive results that have been attributed to genetic diversity of HIV, but really they could reflect other facts that we present in the next section.
Could have HIV sequences an endogenous origin?

When HIV was described, almost nobody knew anything about endogenous retroviruses, although they had been described in animals. Based on not well designed epidemiological surveys, most researchers attributed an exogenous origin to HIV. In 1990s, and especially from 2001, when the first draft of human genome sequence was published, scientific community was aware that 8-10% of human genome consisted of endogenous retroviral sequences. At first, scientists said that these sequences were residues from evolution, as a fossil. They said that they can’t yield mature viral particles, but currently we know there are several endogenous retroviral families that can yield viral particles with infective ability. Moreover, we currently know that these sequences act during intrauterine development to form the placenta, and perhaps influence the temporary immunosupression in pregnant women that allows the fetus development without rejection. Some scientists say that these particles could be a mean of intercellular communication when a stressful event is occurring, in order to improve the adaptation of cells to the new situation.

Related to HIV, in 1992 some endogenous sequences similar to a fragment of HIV was reporting, and other reports exist about the presence of sequences similar to HIV env gene in patients and healthy relatives with prostate and breast cancer. In the last years, elevated viral load of some other endogenous retroviral families not similar to HIV has been described in HIV/AIDS subjects, but not in healthy subjects.

One feature of endogenous sequences is the transmission within chromosomes from parents to children, as other human genes. If so, DNA of these sequences should be present in one or both parents of a HIV+ subject. According to our hypothesis, only when a stressful event occurs, these sequences can be activated, and viral load is detectable, what means that this RNA is translated in proteins that yield a viral particle, with a biological function. Our immune system would produce antibodies against them in order to limit and control their activity. Hence, only persons that were carriers of these sequences could be HIV+ if these ones were activated in stressful conditions, such as drug abuse, exposition to bacterial challenges and injuries during anal intercourses, or blood transfusions.

As previously mentioned, if this hypothesis is true, DNA HIV PCR should be positive in one or both parents of a HIV+ subject. We propose here the following research. From samples belonging to several HIV+ individuals with both parents alive, and some brother or daughter, PCR can be used to detect HIV DNA sequences, by employing primer sequences previously used in commercial kits, such as Roche’s or Abbott’s, and preparing home-made chemical mixtures in order to control the presence of interfering elements. Agarose electrophoresis should be a good procedure for visualization of only one specific band or several bands, and subsequently a DNA sequencing step should be critical for confirming that amplified products belongs to previously described HIV sequences. If this was possible in HIV+ subjects, then the procedure would be performed with samples from relatives.
We have carried out a pilot study in this sense, by using primers used in real time PCR Abbott kits and conditions according to published papers and patents, with samples from a HIV+ woman and her HIV- mother. In both cases, by using different temperature profiles, we obtained a lot of electrophoretical bands that made really impossible to conclude what we are truly amplifying. Although this experiment is preliminar and we cannot extract definitive conclusions, we can suspect that is extremely difficult to consider as reliable those results usually obtained with commercial kits. A more exhaustive research improving PCR conditions (without a quantitative purpose, only for detecting DNA sequences) could give results about the endogenous origin of those HIV sequences previously considered as exogenous.

Juan Manuel Morillo Velázquez

- Bachelor in Dentistry, 1994
- PhD, 2002
- Several postgraduate courses in molecular diagnosis, HIV/AIDS and drug abuse.
- Research experience in application of PCR to microbial and genetic diagnosis of oral conditions, with several papers published in international peer-reviewed journals.
- Clinical experience with subjects diagnosed as HIV+ since 1997, firstly according to official theory of HIV/AIDS, but with a closer viewpoint to dissident positions since 2004.