We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

5,300 Open access books available
131,000 International authors and editors
155M Downloads

154 Countries delivered to
TOP 1% Our authors are among the most cited scientists
12.2% Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Chapter

Polymerase Chain Reaction (PCR): Principle and Applications

Karim Kadri

Abstract

The characterization of the diversity of species living within ecosystems is of major scientific interest to understand the functioning of these ecosystems. It is also becoming a societal issue since it is necessary to implement the conservation or even the restoration of biodiversity. Historically, species have been described and characterized on the basis of morphological criteria, which are closely linked by environmental conditions or which find their limits especially in groups where they are difficult to access, as is the case for many species of microorganisms. The need to understand the molecular mechanisms in species has made the PCR an indispensable tool for understanding the functioning of these biological systems. A number of markers are now available to detect nuclear DNA polymorphisms. In genetic diversity studies, the most frequently used markers are microsatellites. The study of biological complexity is a new frontier that requires high-throughput molecular technology, high speed computer memory, new approaches to data analysis, and the integration of interdisciplinary skills.

Keywords: PCR, molecular markers, genetic diversity

1. Introduction

Polymerase chain reaction (PCR) was invented by Mullis in 1983 and patented in 1985. Its principle is based on the use of DNA polymerase which is an in vitro replication of specific DNA sequences. This method can generate tens of billions of copies of a particular DNA fragment (the sequence of interest, DNA of interest, or target DNA) from a DNA extract (DNA template). Indeed, if the sequence of interest is present in the DNA extract, it is possible to selectively replicate it (we speak of amplification) in very large numbers. The power of PCR is based on the fact that the amount of matrix DNA is not, in theory, a limiting factor. We can therefore amplify nucleotide sequences from infinitesimal amounts of DNA extract. PCR is therefore a technique of purification or cloning. DNA extracted from an organism or sample containing DNAs of various origins is not directly analyzable. It contains many mass of nucleotide sequences. It is therefore necessary to isolate and purify the sequence or sequences that are of interest, whether it is the sequence of a gene or noncoding sequences (introns, transposons, mini or microsatellites). From such a mass of sequences that constitutes the matrix DNA, the PCR can therefore select one or more sequences and amplify them by replication to tens of billions of copies. Once the reaction is complete, the amount of matrix DNA that is not in the area of interest will not have varied. In contrast, the amount of the amplified sequence(s)
(the DNA of interest) will be very big. PCR makes it possible to amplify a signal from a background noise, so it is a molecular cloning method, and clone comes back to purity.

There are many applications of PCR. It is a technique now essential in cellular and molecular biology. It permits, especially in a few hours, the “acellular cloning” of a DNA fragment through an automated system, which usually takes several days with standard techniques of molecular cloning. On the other hand, PCR is widely used for diagnostic purposes to detect the presence of a specific DNA sequence of this or that organism in a biological fluid. It is also used to make genetic fingerprints, whether it is the genetic identification of a person in the context of a judicial inquiry, or the identification of animal varieties, plant, or microbial for food quality testing, diagnostics, or varietal selection. PCR is still essential for performing sequencing or site-directed mutagenesis. Finally, there are variants of PCR such as real-time PCR, competitive PCR, PCR in situ, RT-PCR, etc.

At present, the revolutionary evolutions of the molecular biological research are based on the PCR technique which provides the suitable and specific products especially in the field of the characterization and the conservation of the genetic diversity. Several applications are possible in downstream of the PCR technique: (1) the establishment of a complete sequence of the genome of the most important livestock breeds; (2) development of a technology measuring scattered polymorphisms at loci throughout the genome (e.g., SNP detection methods); and (3) the development of a microarray technology to measure gene transcription on a large scale. The study of biological complexity is a new frontier that requires high throughput molecular technology, high speed and computer memory, new approaches to data analysis, and the integration of interdisciplinary skills.

2. Principle of the PCR

PCR makes it possible to obtain, by in vitro replication, multiple copies of a DNA fragment from an extract. Matrix DNA can be genomic DNA as well as complementary DNA obtained by RT-PCR from a messenger RNA extract (poly-A RNA), or even mitochondrial DNA. It is a technique for obtaining large amounts of a specific DNA sequence from a DNA sample. This amplification is based on the replication of a double-stranded DNA template. It is broken down into three phases: a denaturation phase, a hybridization phase with primers, and an elongation phase. The products of each synthesis step serve as a template for the following steps, thus exponential amplification is achieved [1].

The polymerase chain reaction is carried out in a reaction mixture which comprises the DNA extract (template DNA), Taq polymerase, the primers, and the four deoxyribonucleoside triphosphates (dNTPs) in excess in a buffer solution. The tubes containing the mixture reaction are subjected to repetitive temperature cycles several tens of times in the heating block of a thermal cycler (apparatus which has an enclosure where the sample tubes are deposited and in which the temperature can vary, very quickly and precisely, from 0 to 100°C by Peltier effect) [1, 2]. The apparatus allows the programming of the duration and the succession of the cycles of temperature steps. Each cycle includes three periods of a few tens of seconds. The process of the PCR is subdivided into three stages as follows:

2.1 The denaturation

It is the separation of the two strands of DNA, obtained by raising the temperature. The first period is carried out at a temperature of 94°C, called the denaturation
Polymerase Chain Reaction (PCR): Principle and Applications
DOI: http://dx.doi.org/10.5772/intechopen.86491

temperature. At this temperature, the matrix DNA, which serves as matrix during
the replication, is denatured; the hydrogen bonds cannot be maintained at a tem-
perature higher than 80°C and the double-stranded DNA is denatured into single-
stranded DNA (single-stranded DNA).

2.2 Hybridization

The second step is hybridization. It is carried out at a temperature generally
between 40 and 70°C, called primer hybridization temperature. Decreasing the
temperature allows the hydrogen bonds to reform and thus the complementary
strands to hybridize. The primers, short single-strand sequences complementary to
regions that flank the DNA to be amplified, hybridize more easily than long strand
matrix DNA. The higher the hybridization temperature, the more selective the
hybridization, the more specific it is.

2.3 Elongation

The third period is carried out at a temperature of 72°C, called elongation
temperature. It is the synthesis of the complementary strand. At 72°C, Taq poly-
merase binds to primed single-stranded DNAs and catalyzes replication using the
deoxyribonucleoside triphosphates present in the reaction mixture. The regions of
the template DNA downstream of the primers are thus selectively synthesized. In
the next cycle, the fragments synthesized in the previous cycle are in turn matrix
and after a few cycles, the predominant species corresponds to the DNA sequence
between the regions where the primers hybridize. It takes 20–40 cycles to synthe-
size an analyzable amount of DNA (about 0.1 μg). Each cycle theoretically doubles
the amount of DNA present in the previous cycle. It is recommended to add a final
cycle of elongation at 72°C, especially when the sequence of interest is large (greater
than 1 kilobase), at a rate of 2 minutes per kilobase [1–3]. PCR makes it possible to
amplify sequences whose size is less than 6 kilobases. The PCR reaction is extremely
rapid, it lasts only a few hours (2–3 hours for a PCR of 30 cycles).

2.4 Primers

To achieve selective amplification of nucleotide sequences from a DNA extract
by PCR, it is essential to have least one pair of oligonucleotides. These oligonucle-
otides, which will serve as primers for replication, are synthesized chemically
and must be the best possible complementarity with both ends of the sequence of
interest that one wishes to amplify. One of the primers is designed to recognize
complementarily a sequence located upstream of the fragment 5′–3′ strand DNA
of interest; the other to recognize, always by complementarity, a sequence located
upstream complementary strand (3′–5′) of the same fragment DNA. Primers are
single-stranded DNAs whose hybridization on sequences flanking the sequence
of interest will allow its replication so selective. The size of the primers is usually
between 10 and 30 nucleotides in order to guarantee a sufficiently specific hybrid-
ization on the sequences of interest of the matrix DNA [1–5].

2.5 Taq polymerase

DNA polymerase allows replication. We use a DNA polymerase purified or
cloned from of an extremophilic bacterium, Thermus aquaticus, which lives in hot
springs and resists temperatures above 100°C. This polymerase (Taq polymerase)
has the characteristic remarkable to withstand temperatures of around 100°C,
which are usually sufficient to denature most proteins. *Thermus aquaticus* finds its temperature of comfort at 72°C, optimum temperature for the activity of its polymerase [4].

3. The reaction conditions

The volumes of reaction medium vary between 10 and 100 μl. There are a multitude of reaction medium formulas. However, it is possible to define a standard formula that is suitable for most polymerization reactions. This formula has been chosen by most manufacturers and suppliers, who, moreover, deliver a ready-to-use buffer solution with Taq polymerase. Concentrated 10 times, its formula is approximately the following: 100 mM Tris-HCl, pH 9.0; 15 mM MgCl₂, 500 mM KCl [2, 4].

It is possible to add detergents (Tween 20, Triton X-100) or glycerol in order to increase the conditions of stringency that make it harder and therefore more selective hybridization of the primers. This approach is generally used to reduce the level of nonspecific amplifications due to the hybridization of the primers on sequences without relationship with the sequence of interest. We can also reduce the concentration of KCl until eliminated or increase the concentration of MgCl₂ [1, 5]. Indeed, some pairs of primers work better with solutions enriched with magnesium. On the other hand, with high concentrations of dNTP, the concentration of magnesium should be increased because of stoichiometric interactions between magnesium and dNTPs that reduce the amount of free magnesium in the reaction medium. dNTPs (deoxyribonucleoside triphosphates) provide both the energy and the nucleotides needed for DNA synthesis during the chain polymerization. They are incorporated in the reaction medium in excess, that is, about 200 μM final. Depending on the reaction volume chosen, the primer concentration may vary between 10 and 50 pmol per sample. Matrix DNA can come from any organism and even complex biological materials that include DNAs from different organisms. But to ensure the success of a PCR, it is still necessary that the DNA matrix is not too degraded. This criterion is obviously all the more crucial as the size of the sequence of interest is large. It is also important that the DNA extract is not contaminated with inhibitors of the polymerase chain reaction (detergents, EDTA, phenol, proteins, etc.) [6, 7]. The amount of template DNA in the reaction medium initiates that the amplification reaction can be reduced to a single copy. The maximum quantity may in no case exceed 2 μg. In general, the amounts used are in the range of 10–500 ng of template DNA. The amount of Taq polymerase per sample is generally between 1 and 3 units. The choice of the duration of the temperature cycles and the number of cycles depends on the size of the sequence of interest as well as the size and the complementarity of the primers. The durations should be reduced to a minimum not only to save time but also to prevent risk of nonspecific amplification. For denaturation and hybridization of primers, 30 seconds are usually sufficient. For elongation, it takes 1 minute per kilobase of DNA of interest and 2 minutes per kilobase for the final cycle of elongation. The number of cycles, generally between 20 and 40, is inversely proportional to the abundance of DNA matrix [6–8].

4. PCR product detection and analysis

The product of a PCR consists of one or more DNA fragments (the sequence or sequences of interest). The detection and analysis of the products can be very quickly carried out by agarose gel electrophoresis (or acrylamide). The DNA is revealed by ethidium bromide staining [2, 3, 5]. Thus, the products are instantly
visible by ultraviolet transillumination (280–320 nm). Very small products are often visible very close to the migration front in the form of more or less diffuse bands. They correspond to primer dimers and sometimes to the primers themselves. Depending on the reaction conditions, nonspecific DNA fragments may be amplified to a greater or lesser extent, forming net bands or “smear” [6–9]. On automated systems, a fragment analyzer is now used. This apparatus uses the principle of capillary electrophoresis. Fragment detection is performed by a laser diode. This is only possible if the PCR is performed with primers coupled to fluorochromes [10].

5. Overview of molecular techniques based on PCR technology

Microsatellites are hypervariable; on a locus, they often show dozens of different alleles from each other in the number of repetitions. They are still the markers of choice for studies on the diversity, paternity analysis and mapping of quantitative effects loci (QTL), although this could change, in the near future, through the elaboration inexpensive SNP assay methods. Minisatellites have the same characteristics as microsatellites, but repetitions range from ten to a few hundred base pairs. Micro- and minisatellites are also known as variable number of tandem repeat (VNTR) polymorphisms. Amplified fragment length polymorphisms.

5.1 Microsatellites

Microsatellites are now the most used markers in genetic characterization studies of farmed animals [11]. The high mutation rate and codominant nature favor the estimation of intra and interracial diversity, and the genetic mixing between races, even if they are very close. Challenges have surrounded the choice of a mutation model—the infinite or progressive allele mutation model [12] for the analysis of microsatellite data. However, simulation studies have indicated that the infinite allele mutation model is generally valid for the evaluation of intraracial diversity [13]. The low number of alleles per population and observed and expected heterozygosity are the most commonly used parameters for assessing intraracial diversity. The simplest parameters for evaluating interracial diversity are genetic differentiation or fixation indices. Several estimators have been proposed (e.g., FST—fixation index and GST—glutathione S transferase), and the most widely used is FST [14], which measures the degree of genetic differentiation of subpopulations by calculation standardized variances of allele frequencies of populations. Statistical significance is calculated for FST values between population pairs to test the null hypothesis of a lack of genetic differentiation between populations and, consequently, the partitioning of genetic diversity [15]. Microsatellite data are also commonly used to assess genetic relationships between populations and subjects through the estimation of genetic distances [16–19]. The measure of genetic distance used most often is the standard genetic distance of Nei [20]. In another case, the modified Cavalli-Sforza distance is recommended [21] for the closest populations, where genetic drift is the main factor of genetic differentiation. The genetic relationship between breeds is often visualized by the reconstruction of a phylogeny, most often using the “neighbor-joining” method [22]. However, the main problem in the reconstruction of the phylogenetic tree is that the evolution of the lines is presumed to be uncrosslinked that is to say that the lines can deviate, but can never come from interbreeding. This assumption is rarely valid for farm animals, as new breeds are often derived from crosses between two or more ancestral breeds. The visualization of breed evolution by phylogenetic reconstruction must therefore be interpreted with great attention.
5.2 SNP

Single nucleotide polymorphisms (SNPs) are used as an alternative to microsatellites in genetic diversity studies. Several technologies are available to detect the type of SNP markers [23]. As biallelic markers, SNPs have relatively low amounts of information, and to reach the information level of a standard panel of 30 microsatellite loci, larger amounts must be used. However, ever-evolving molecular technologies increase automation and reduce the cost of typing SNPs, which will likely allow, in the near future, the parallel analysis of a large number of markers at a reduced cost. In this perspective, large-scale projects are being implemented for several livestock species to identify millions of SNPs [24] and validate several thousands and identify haplotype in the genome. As with sequence information, SNPs allow for direct comparison and joint analysis of different experiments. SNPs are likely to be interesting markers for future use in genetic diversity studies because they can be easily used in the assessment of functional or neutral variation. However, the preliminary phase of SNP discovery or selection of SNPs from databases is critical. SNPs can be generated through various experimental protocols, such as sequencing, single-stranded coformational polymorphism (SSCP) or denaturing high-performance liquid chromatography (DHPLC) or in silico, aligning and comparing multiple sequences from the same region from public databases on genomes and sequential expression tags (ESTs). If the data were obtained randomly, the standard population genetic parameter estimators cannot be applied. A common example is when SNPs initially identified in a small sample (panel) of individuals are then typed into a larger sample of chromosomes. By preferably performing sampling of SNPs at intermediate frequencies, such a protocol will affect the distribution of allele frequencies with respect to the probable values for a random sample. SNPs present a modern tool in the context of genetic analyzes of the population; however, it is necessary to develop statistical methods that will take into account each SNP operating method and their locations [25, 26].

5.3 Amplification of fragment length polymorphism (AFLP)

AFLPs are dominant biallelic markers [27]. Variations on many loci can be arranged simultaneously to detect single nucleotide variations of unknown genomic regions, where a given mutation may often be present in undetermined functional genes. The disadvantage is that they show a dominant mode of inheritance, which reduces their power during genetic analyses of the population on intraracial diversity and consanguinity. However, AFLP profiles are highly informative in the evaluation of race relations [28–32] and related species [33].

5.4 Restriction fragment length polymorphism (RFLP)

Restriction fragment length polymorphisms (RFLPs) are identified using restriction enzymes that cut DNA only at specific “restriction sites” (e.g., EcoRI cuts at the site defined by the palindrome GAATTC sequence). At present, the most common use of RFLPs is downstream PCR (PCR-RFLP) to detect alleles that differ in sequence at a given restriction site. A gene fragment is first amplified using PCR and then exposed to a specific restriction enzyme that cuts only one of the allelic forms. The digested amplicons are usually resolved by electrophoresis. Microsatellites or SSRs (simple sequence repeats) or STRs (short tandem repeats) consist of a few nucleotides—2–6 base pair DNA sequence—repeated several times in tandem (e.g., CACACACACACACACA). They are spread on a eukaryotic genome. Microsatellites are relatively small in size and, therefore, are easily
amplified using DNA PCRs extracted from different sources, such as blood, hair, skin, or even feces. Polymorphisms can be visualized on a sequencing gel, and the availability of automated DNA sequencers allows high-throughput analysis of a large number of samples [34, 35].

5.5 Mitochondrial DNA markers

Mitochondrial DNA polymorphisms (mtDNA) have been widely used in analyzes of phylogenetic and genetic diversity. The haploid mtDNA transported by the mitochondria of the cellular cytoplasm has a maternal mode of inheritance (the animals inherit the mtDNA from their mothers and not from their fathers) and a high mutation rate; it does not recombine. These features allow biologists to reconstruct intra and interracial evolutionary relationships by evaluating mtDNA mutation patterns. mtDNA tags can also provide a quick way to detect hybridization between farmed species and subspecies [36]. Polymorphisms in the hypervariable region of the D-loop or the mtDNA control region have largely contributed to the identification of wild ascendants of domestic species and to the establishment of geographical models of genetic diversity.

6. Applications

6.1 Acellular cloning

This is one of the most remarkable applications of PCR. It makes it possible to isolate, that is to say, to purify a gene without resorting to traditional methods of molecular cloning which consist in inserting a DNA library in a plasmid vector which is then used to transform a bacterial strain whose clones after selection are screened. The realization is much faster and much less random using PCR. Acellular cloning is used when using PCR because it is useless to use a cellular system (bacteria, yeast, and animal or plant cell) to amplify the clone. The realization of molecular cloning by PCR depends on two major criteria: the choice of DNA extract (matrix DNA) and primers. It is indeed essential to have more or less reliable data on the sequence of the gene that is to be cloned and/or flanking sequences in order to synthesize the sets of primers necessary for its amplification in whole or in part. On the other hand, it is still necessary to perform the PCR on the appropriate matrix DNA [37, 38]. We can choose the genomic DNA that includes the total sequence of the genome and therefore all the genes of the species. In this case, the genes include both exons and introns and their amplification results in the cloning of the complete gene sequence and even, depending on the primers that have been chosen, regulatory regions. But we can also choose to extract the messenger RNA (mRNA), that is to say the only coding sequences of the gene—the transcripts. Since RNAs are unstable, messenger RNAs are transformed into complementary DNA (cDNA) by RT-PCR (see below), a variant of PCR that uses reverse transcriptase and allows changing the RNA sequences into DNA. It is on this cDNA library that PCR is then performed to clone the gene of interest. In this case, the deal is more complex. The presence of the gene transcript in the extract depends on the cell type, tissue, or organ from which the mRNA extraction was performed. Indeed, transcription is specific to the cell type. More serious, the expression of a gene is often regulated by physiological factors, environmental, in this case the gene of interest is not necessarily transcribed and the cDNA library may not contain it. Finally, it must be said that transcription is itself regulated and is often accompanied by alternative splicing. This phenomenon leads to exon elimination at the time of excision of the
introns and leads to the expression of different proteins from the same gene. It follows that depending on the cell type and regulatory profiles, we may not be dealing with the same transcript. It is nevertheless very interesting to clone a transcript since its nucleotide sequence corresponds to the amino acid sequence resulting from the translation. On the other hand, with a cDNA, it is easier to carry out the expression of the gene and thus the functional evaluation of the corresponding protein or proteins in a cellular model of expression. Very frequently, PCR cloning is practiced in parallel on genomic DNA (genomic library) and different cDNA libraries so as to determine the complete sequence of the gene, its expression profile, the modalities of splice regulation [8, 39], etc.

6.2 Reverse transcriptase PCR (RT-PCR)

As discussed in the previous chapter, it may be relevant to extract the mRNAs to then generate cDNA copies. This reaction is catalyzed by retrovirus reverse transcriptase (reverse transcriptase) which synthesizes a DNA chain from an RNA template. At first, the total RNAs are extracted. The mRNAs are isolated from the total RNA by affinity chromatography using oligoT (polyT oligonucleotide) because the messenger RNAs are characterized by a 3' polyA sequence. Then, the mRNAs are subjected to reverse transcriptase which will generate a copy of DNA (cDNA) of each mRNA. After the reverse transcription, the mRNAs are hydrolyzed (alkaline treatment, RNase, or temperature). The following steps are carried out in the enclosure of the thermal cycler. The single-stranded cDNAs are then replicated by the DNA polymerase during a first temperature cycle [40, 41]. Other cycles are repeated to amplify double-stranded cDNAs in large quantities. In a given cell phenotype, an estimated 10–15,000 genes are expressed in humans and most mammals. Some cell transcripts are expressed at a few hundred or even a few thousand copies per cell, but the majority of transcripts represent a low copy number. The expression profiles of transcripts undergo qualitative or quantitative variations that reflect the biological dynamics of the cell. The identification of variations in gene expression in a given physiological or pathological context can therefore provide valuable information concerning the function of genes and the influence of modulation factors on their expression, whether they are physiological or of environmental origin. The analysis of the expression variations of genes involved in a pathology can lead to new therapeutic or diagnostic targets. Finally, from a fundamental point of view, studying the gene expression profile makes it possible to advance in understanding the mechanisms of cellular physiology [40–42].

6.3 Quantitative PCR in real time (quantitative real-time PCR)

Developed in the mid-1980s, quantitative PCR can determine the level of specific DNA or RNA in a biological sample. The method is based on the detection of a fluorescent signal that is produced in proportion to the amplification of the PCR product, cycle after cycle. It requires a thermal cycler coupled to an optical reading system that measures fluorescence emission. A nucleotide probe is synthesized so that it can hybridize selectively to the DNA of interest between the sequences where the primers hybridize. The probe is labeled on the 5' end with a fluorochrome signal (e.g., 6-carboxyfluorescein), and on the 3' end with a quencher (e.g., 6-carboxytetramethyl rhodamine). This probe must show temperature hybridization (Tm) greater than that of the primers so that it hybridizes 100% during the elongation phase (critical parameter) [43–45].

As long as the two fluorochromes remain present at the probe, the extinguisher prevents the fluorescence of the signal. In this step, the proximity of the quencher
and the signal induces a lack of fluorescence emission. During this phase of elongation, Taq polymerase, which has an intrinsic 5′–3′ nuclease activity, degrades the probe and thus releases the fluorochrome signal. The level of fluorescence then released is proportional to the amount of PCR products generated in each cycle. The thermal cycler is designed so that each sample (the PCR is generally carried out in 96-well plates) is connected to an optical system. This includes a laser transmitter connected to an optical fiber. The laser, via the optical fiber, excites the fluorochrome within the PCR reaction mixture. The fluorescence emitted is retransmitted, always through optical fiber, to a digital camera connected to a computer. A software then analyzes and stores the data. Quantitative PCR is a method of high specificity and sensitivity. It is very timely for countless applications. A conventional PCR only provides qualitative data (presence or absence of the DNA of interest, purification of this DNA). Quantitative PCR, as its name suggests, makes it possible to know more precisely the quantity of the DNA of interest (or RNA, since it is possible to conduct a quantitative RT-PCR with the same apparatus) [45–47]. It is indeed very often used for this purpose, in order to determine the viral load, in particular in cases of hepatitis C or AIDS. One of the most remarkable and useful applications is the analysis of gene expression through the quantitative measurement of transcripts.

6.4 Semi-quantitative or competitive PCR

This is in most cases RT-PCR. In the case of quantitative PCR, the level of RNA or DNA of interest is measured as the absolute amount. In the case of semi-quantitative PCR or competitive PCR, it is a question of measuring relative quantities by means of standards that correspond to RNA or more rarely to DNA. This is in most cases RT-PCR. These standards can be internal or external. External standards may be homologous or heterologous. The standard is an RNA (more rarely a DNA) which is present in the RNA extract (internal standard) or which is added in known quantity in the reaction mixture (external standard). The standard is amplified at the same time as the RNA of interest. There is therefore competition between the amplification of the standard and that of the DNA of interest. The higher the standard quantity, the less the RNA of interest will be amplified and therefore its quantity will be small. Of course, the method of analysis of the PCR sample must make it possible to discriminate the standard with respect to the RNA of interest on the one hand and on the other hand to evaluate the relative amount of DNA of interest by comparison with the amount of standard that is known [48]. The internal standards are endogenous RNA, corresponding to RNA genes whose expression is presumed constant (actin, beta2-microglobulin, etc.) and which are present in the population of RNA matrices during reverse transcription. These standards have a major disadvantage: they require the use of primers different from those used for the RNA of interest. The kinetics of amplification are therefore substantially different, and it is very difficult or impossible to guarantee a constant expression between different samples. The homologous external RNA standards are synthetic RNAs that share the same priming hybridization sites as the RNA of interest and that have the same overall sequence, with a slight mutation, deletion, or insertion that will allow the identification and quantification thereof with respect to the signal rendered by the RNA of interest. These standards make it possible on the one hand to appreciate the variability introduced at the level of the RT and, on the other hand, generally have the same amplification efficiency as the RNA of interest whether it is at the RT level or PCR [48, 49].

The heterologous external RNA standards are exogenous RNAs and their rate can therefore be controlled. However, unlike homologous external standards, they
have a different amplification efficiency compared to that of the RNA of interest. In the case of quantitative RT-PCR (semi-quantitative PCR), the standard consists of a titrated solution of DNA of sequence identical to that of the DNA of interest to be quantified. A dilution series is performed, each being used for amplification. It is then a question of defining the ideal number of cycles to be placed in the exponential phase of the reaction while ensuring an effective amplification. Then, each standard DNA dilution as well as the DNA extracted from the sample to be quantified are submitted in parallel to the PCR reaction. A standard curve is established with standard dilutions \([\text{signal} = f(\text{concentration})]\). Knowing the value of the signal measured on the sample to be quantified, the corresponding number of copies can be extrapolated from the curve. In the case of competitive PCR, a series of synthetic external homologous standard RNA dilutions are co-amplified with equivalent amounts of total RNA (and thus an equivalent amount of the native gene) \([50, 51]\). The standard competes with the RNA of interest for polymerase and primers. As the standard concentration increases, the signal of the gene of interest decreases. Here, the PCR does not need to be performed in the exponential phase and the results show a correct reproducibility. However, the method is cumbersome and does not allow to manage many samples simultaneously \([52]\).

### 6.5 PCR applied to diagnosis

PCR is a fabulous diagnostic tool. It is already widely used in the detection of genetic diseases. The amplification of all or part of a gene responsible for a genetic disease makes it possible to reveal the deleterious mutations (s), their positions, their sizes, and their natures. It is thus possible to detect deletions, inversions, insertions, and even point mutations, either by direct analysis of PCR products by electrophoresis or by combining PCR with other techniques \([53]\). But PCR can still be used to detect infectious diseases (viral, bacterial, parasitic, etc.), as is already the case for AIDS, hepatitis C, or chlamydia infections. Although other diagnostic tools are effective at detecting these diseases, PCR has the enormous advantage of producing very reliable and rapid results from minute biological samples in which the presence of the pathogen is not always detectable with other techniques \([53, 54]\).

### 6.6 Detection of genetic diseases

In the context of genetic diseases, it is a question of detecting a mutation on the sequence of a gene. Several situations arise. The simplest ones concern insertions and deletions. In these cases, the mutation is manifested by the change in the size of the gene or part of the gene. Insofar as the mutation is known and described, it suffices to amplify all or part of the gene. In the case of an insertion, the PCR product from a patient’s DNA is longer than that from a healthy person. A deletion presents a contrary result \([55]\). The analysis of PCR products by electrophoresis, and therefore the evaluation of their size, leads directly to the diagnosis. The detection of inversions and point mutations is more delicate. The difference in size between healthy and diseased DNA is zero in the case of an inversion and almost zero in the case of a point mutation. We cannot therefore retain the size criterion of the PCR products to achieve the result. It is therefore necessary to resort to techniques complementary to PCR. Three approaches can be selected, the southern blot, the restriction fragment length polymorphism (RFLP), or the detection of mismatch. The southern blot consists in hybridizing on the PCR product an oligonucleotide probe marked, thanks to a radioactive isotope or a fluorochrome, whose sequence is complementary and therefore specific to that which corresponds to the mutation. This strategy is well suited to inversion cases \([56, 57]\).
The RFLP can detect inversions such as point mutations. It involves a restriction enzyme capable of hydrolyzing the PCR product at the sequence which sets the mutation. This approach is only possible if a restriction site is indeed present on this sequence, whether it is the mutated allele or the wild-type allele. The restriction enzyme thus hydrolyzes either the PCR product derived from healthy DNA or that which is derived from the diseased DNA. From these PCR products, one or two DNA fragments are thus obtained which are then revealed by electrophoresis. Mismatch detection is, like the RFLP, adapted to inversions and point mutations [57–59]. The PCR product from the patient’s DNA (sample DNA) is mixed with the PCR product from the DNA of a healthy person (reference DNA). This mixture is then denatured by the temperature and then rehybridized. Yes, the sample DNA is mutated; the pairings between sample DNA and reference DNA will be incomplete at the level of the mutation. The mismatches concern a single base pair in the case of a point mutation and several base pairs in the case of an inversion. These mismatches are then degraded by S1 nuclease, an enzyme that degrades only single-stranded DNAs. Another solution is to cleave the mismatches chemically (osmium tetroxide, then piperidine), but it is more suitable for point mutations. In summary, mutation induces a mismatch at the level of enzymatic or chemical cleavage which leads to the generation of two fragments from a single PCR product. These fragments are analyzed by electrophoresis.

6.7 Detection of infectious diseases

Contamination with viruses or microorganisms (bacteria, parasites, etc.) necessarily results in the presence of their genetic material in all or part of the infected organism. PCR is therefore a tool all the more effective in detecting the presence of a pathogen in a biological sample that its sensitivity and specificity are very large. The performance of the PCR diagnosis is essentially based on a criterion: the choice of primers capable of very selectively amplifying a sequence of the DNA of the virus or microorganism [57–59]. Matrix DNA, on the other hand, must be extracted from a tissue in which the microorganism is present. It is therefore sufficient to amplify a specific sequence of the pathogen from a sample taken on the patient and to analyze the PCR product by electrophoresis. The size of the amplified DNA fragment, which must conform to the expected size, guarantees the reliability of the result and therefore of the diagnosis. In the case of AIDS (HIV) testing, for example, routine testing is based on the ELISA method of detecting HIV antibodies or viral antigens in the patient’s serum by an immunoassay technique. This method, quite reliable and inexpensive, nevertheless has some disadvantages. False positives are quite common because of cross-reactivities. Positive samples are therefore tested for control by another routine technique, Western blot. There remains the problem of HIV-positive people who do not carry the virus, such as children whose mothers have AIDS. The blood of these newborns usually contains anti-HIV antibodies of maternal origin and they are therefore seropositive. On the other hand, they do not necessarily carry the virus. In this type of case, the PCR diagnosis is relevant [57–60]. The method involves amplifying a specific sequence of the provirus from a lymphocyte extract. The same principle is used for the detection of toxoplasma in newborns whose mother is a carrier. It is of course possible to diagnose AIDS by RT-PCR by looking for viral RNA in the patient’s serum. Quantitative or semi-quantitative methods have been developed which also make it possible to evaluate the viral load.

6.8 PCR applied to identification

PCR is remarkably effective at identifying species, varieties, or individuals by genetic fingerprinting. This application is based on the knowledge acquired on
genome structure. It is simply to amplify nucleotide sequences that are specific to species, variety, or individual. In eukaryotes, in particular, these sequences are very numerous and offer a vast palette that allows identification in a very precise and very selective way. Indeed, the genomes of eukaryotic organisms have, unlike prokaryotes, coding sequences and noncoding sequences. The coding sequences correspond to the genes and are therefore translated into proteins. The noncoding sequences, which are therefore not translated, represent a large proportion of eukaryotic genomic DNA (up to 98%). The coding sequences are highly homologous in individuals of the same species. Indeed, the species is characterized by characters and common traits that are guaranteed by its genes. The phenotypic differences between the individuals that compose it are based on the allelic variations and the different alleles of the same gene show sequence differences that are minute (of the order of 1 base pair per 1000) [61, 62]. From one species to another, depending on the phylogenetic distance that separates them, the sequences of the genes that code for the same function have very strong homologies, all the more so that the function of the gene is essential to the embryogenesis or metabolism. As a result, coding sequences are of little relevance in terms of identification. On the other hand, the noncoding sequences are very polymorphous between species as between individuals of the same species. They thus present a large choice of genetic markers that make it possible to establish identification tests which are highly discriminating. Among these markers are minisatellites (or variable number of tandem repeats) and microsatellites (or STR, short tandem repeats) [61–63]. VNTRs and STRs are repetitive polymorphisms composed of sequences that are repeated in tandem. These repeat sequences measure from 10 to 40 base pairs for VNTRs and from 1 to 5 base pairs for STRs. From one individual to another, the repeated sequence of a VNTR or STR is identical but the number of repetitions and therefore the size of the VNTR or the STR can be very variable (we speak of alleles). On the other hand, there is a wide variety of VNTRs and STRs on eukaryotic genomes. Detection of STR or VNTR polymorphism is by PCR using primers that hybridize to nonpolymorphic flanking sequences. The amplification products are then either analyzed by electrophoresis or undergo fragment analysis using a capillary sequencer. It is now possible to simultaneously amplify several STRs or VNTRs by using several pairs of primers. The variety of amplification products obtained leads to footprints that are specific individuals. On the other hand, the power of PCR makes it possible to amplify micro- and minisatellites from very little DNA. DNA fingerprinting has become much more commonplace in recent years in the context of judicial investigations. But these techniques are equally as effective in other species as humans and allow not only identifying individuals but also varieties or species. The type of identification depends simply on the choice of markers. Similarly, for varietal identification purposes, one can commonly proceed according to protocols derived from the PCR [64–66].

Two techniques that are relevant are the random amplification of polymorphic DNA (RAPD) and the amplification of fragment length polymorphism (AFLP). (Random amplification of polymorphic DNA (RAPD) is a PCR for varietal identification that uses pairs of random primers of reduced size (about 10 base pairs). These primers will hybridize randomly, but PCR usually results in an electrophoresis amplification profile which is specific to the variety from which the matrix DNA is derived. Amplification of fragment length polymorphism (AFLP) is a much more efficient method. It first consists hydrolyzing the genomic DNA with one or better two restriction endonucleases. Then, we proceed with the ligation of adapters (defined sequences of DNA of about 15 nucleotides) at the level of the generated cohesive ends by restriction enzymes. Finally, the product of the ligation is amplified by PCR with a pair of primers that hybridizes at the level of the adapters. The
AFLP gives a result comparable to the RAPD. However, the AFLP shows cleaner and more reproducible results. This is the most successful method to date applied to varietal identification.

7. Conclusion

The extension of genotyping approaches to all living organisms has made significant advances in the reconstruction of the history of life. At the population level, the distribution and frequency of known genetic polymorphisms in a species can highlight the evolving forces at play, reveal the effects of natural selection, and infer demographic change. Moreover, the comparison of the sequences of the same genes between different species and that of whole genomes is at the origin of the molecular phylogenies that currently prevail in the classification. They make it possible to trace the relationships between species on the basis of the divergence of their DNA sequences. As such, the PCR is a key stage at two levels. The first concerns the isolation of homologous genes in several species and their characterization. The second is the production of amplified total genomic DNA for genome sequencing and comparative analysis. But PCR is also used to identify the genetic heritage of missing organisms. The DNA breaks down by fragmentation after the death of the body. If we can recover these fragments and amplify them, it becomes possible, in spite of its state, to deduce all or part of the initial genome of the individual. PCR has thus become the primary tool in the field of palaeogenetics, which consists in recovering and analyzing DNA sequences of more or less old organisms, and this as well from the remains preserved in museum collections, from historical sites where the skeletal or mummified remains of extinct organisms for hundreds thousands or even hundreds of thousands of years. The uses of the PCR thus quickly stopped being limited to the studies of biology, to gain other disciplines or fields of activities.
References


[25] Nielsen R, Signorovitch J. Correcting for ascertainment biases when analyzing SNP data:


[36] Nijman IJ, Otsen M, Verkaar EL, de Ruijter C, Hanekamp E. Hybridization of banteng (Bos javanicus) and zebu (Bos indicus) revealed by mitochondrial DNA, satellite DNA, AFLP and microsatellites. Heredity. 2003;90:10-16


La RT-PCR en diagnostique clinique. Annale de Biologie Clinique. 2003;61:635-644


[62] Giesendorf BA, Tyagi JA, Mensink S, Trijbergs EF, Blom HJ. Molecular