

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

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

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



Nucleic Acid Isolation and Downstream Applications

Ivo Nikolaev Sirakov

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/61833>

Abstract

Nucleic acids are not only a source of life but also a means of observing, understanding, and regulating it. Nucleic acids, DNA and RNA, and their characteristics are discussed in other chapters of the book. This chapter describes the fundamental principles of different methods for nucleic acid sample preparation / nucleic acid extraction, such as column-based methods using silica membranes and traditional ones without a column purification procedure (commercially available or homemade). Other topics discussed here include comparative analysis of the use of these methods in DNA and RNA extraction from a variety of biological and clinical samples, as well as the relationship between the type of sample, the method used and the quality and amount of extracted DNA or RNA. Finally, the chapter outlines the application of nucleic acids in the diagnosis of various diseases, in scientific research, and bird sex determination by downstream applications such as restriction enzyme analysis, polymerase chain reactions (PCR, reverse transcription-PCR, real-time PCR), and different sequencing methods (Sanger, cycling sequencing, and next-generation sequencing).

Keywords: DNA extraction, RNA extraction, PCR methods, next-generation sequencing

1. Introduction

1.1. Nucleic acid isolation and downstream applications

The specific properties of nucleic acids have been widely employed in the development of different molecular methods and mathematical models for their analysis. These methods are applied to identify microorganisms and genetic predispositions, to detect different mutations and determine their role in antibiotic resistance, to study phylogenetic relationships, and so on. What all these methods share in common is their starting point: obtaining a purified nucleic acid sample.

2. Nucleic acid extraction methods

Since nucleic acid extraction is a starting point in a vast array of downstream applications, the high quality of nucleic acids in the starting samples is a key factor for the success of the subsequent steps of analysis. Thus, nucleic acid extraction could be defined as a series of steps to obtain nucleic acid samples/materials of particular purity that are free of impurities and are suitable for different downstream application steps. The purpose of nucleic acid extraction methods is to disintegrate the cell envelope and achieve maximum elimination of lipids and proteins to obtain pure DNA and/or RNA. This is principally based on heat adsorption on silica membranes/beads, anion exchange chromatography, sedimentation/precipitation, and use of magnetic particles. These methods yield initial nucleic acid samples of different purity and concentration depending on the original sample (bacteria, viruses, tissues).

The choice of method – in view of optimal time/quality balance – depends on the aim of the study, the type of analysis, the type of nucleic acid, and the cost. It is important to provide appropriate conditions for nucleic acid extraction in order to avoid nucleic acid degradation due to oxidation by reactive oxygen species generated during respiration *in vivo* or, extracellularly, by mechanisms involving metal ions [1–4]. Nucleic acid degradation can result from hydrolysis of the 3'-5' phosphodiester bonds catalyzed by metal complexes as well as from the spontaneous breakage of these bonds due to transesterification via a nucleophilic attack at the phosphorus atom by an adjacent 2'-hydroxyl group [1].

What accounts for the differences between the methods for extraction of DNA and RNA is their different stability. RNA includes ribosomal RNA (rRNA) 80%, mitochondrial RNA (mtRNA), messenger RNA (polyadenylated – poly A⁺ in eukaryotic cells) (mRNA) 1–5%, transfer RNA (tRNA), and microRNA molecules (miRNA). There may be different amounts of mRNA in cells: from large quantities to just five copies per cell. In fact, mRNA is the RNA of choice in reverse transcription and cDNA synthesis. RNA molecules are susceptible to degradation by ozone in the air; ozone is highly reactive regardless of whether the RNA sample is liquid or solid [5]. Another factor that plays a role in RNA degradation is water, as it makes the transfer of protons possible and serves as a source of hydronium or hydroxyl ions. That is why dehydration has a protective effect against RNA degradation [6]. Nucleic acids, and especially RNA molecules, are also sensitive to nucleases. Therefore, in RNA extraction procedures, it is essential to ensure an RNase-free fraction and a means to quickly cool down the sample. This illustrates the importance of providing all necessary work facilities: use of BSL-1 or 2 laminar flow cabinets (depending on the type of biological material) is recommended for nucleic acid extraction; ultraviolet (UV) germicidal irradiation of both the premises and the laminar flow cabinets should be done, with irradiation of the premises done the night before (irradiation immediately before work may lead to degradation due to residual UV light). The same principle applies to UV germicidal irradiation of plastic labware, e.g., microtubes, pipette tips, etc. (if not commercially sterile, DNase- and RNase-free). Another important detail is that talc-free gloves should be used because talc may inhibit some downstream analyses such as PCR, reverse transcription, and real-time-PCR.

The quality of the extracted nucleic acids also depends on the quality of the starting sample. In fact, all manufacturers of nucleic acid isolation kits recommend that fresh starting material be used. If this is not possible – as is often the case with diagnostic samples – they can be stored for 24–48 h at 4°C, or for longer periods of time at –80°C or in liquid nitrogen, preferably using protective buffers, especially for samples intended for RNA analysis. There are paper matrices especially developed for storage and transport of blood samples at room temperature – dried blood spot sampling. In this train of thought, it has to be kept in mind that heparin, which is used as an anticoagulant, may inhibit some PCR reactions [7] and should, therefore, be avoided or removed. In the case of clotted blood samples, the coagulum can be treated as an organ sample. The spleen and the liver are transcriptionally active organs and they have a very high RNA content. Because of that, if the samples are intended for DNA analysis, they have to be treated with RNase prior to column purification. When the aim is RNA analysis, it is particularly important to protect the RNA against degradation, else the low-frequency transcripts could be lost and would not be detected in the downstream steps of analysis. Moreover, in microarray analysis, degraded RNA molecules may fail to successfully bind to the complementary site due to loss of the complementary sequence. That is why, in RNA extraction, frozen samples should be mechanically processed (homogenized) prior to thawing, and fresh tissue samples should be ground in liquid nitrogen or by other means of cooling. This is not as essential in DNA extraction, since DNA molecules are relatively more robust, but is recommended.

Another key step is cell lysis, which – if incomplete – would result in reduced yield and column blocking and, in turn, in lower purity. There are different ways to aid the process of cell lysis: in the case of cell cultures or bacterial cultures, depending on the aim of analysis, they can be washed in PBS (phosphate buffered saline) or physiological saline and resuspended in ddH₂O and/or subjected to several freeze/thaw cycles. In the case of mucous samples (nasal discharges, sputum, intestinal loops), it is good to first decrease the viscosity of the material (using a mucolytic –“mucus-dissolving” agent, e.g., acetylcysteine). The mechanical processing of samples from insects, plants, feces, organs requires 50–200 mg of sample in most kits. Gram-positive bacteria are treated with lysozyme; yeasts, with zymolyase or lyticase; and paraffin-embedded tissues are treated with xylene to remove the paraffin. Other approaches that can be applied to disrupt cell envelopes include: osmotic shock, which is suitable for Gram-negative bacteria, cell cultures and erythrocytes; chaotropic salts, for all types of samples with the exception of some Gram-negative bacteria, owing to the greater thickness of their peptidoglycan layer; enzymatic degradation (lysozyme, proteinase K), which is often combined with osmotic shock or freeze/thaw cycles (for DNA extraction from hair, feathers etc.); and detergents, for tissue cultures. Should a sample remain not fully lysed, one way to overcome the problem is to centrifuge the mixture and use the supernatant prior to column loading or alcohol supplementation, in case the protocol includes such a step.

Sometimes the samples may be “old”, i.e., stored for a long time at –20°C. In such samples, the chemical bonds in the DNA molecules may have become weaker. Then, in the vortexing steps in the column-based and solution-based methods, the DNA becomes, more often than not,

degraded, resulting in smeared DNA bands (Figure 1). To avoid this, such samples should be kept frozen during the process of mechanical homogenization, and instead of vortexed, should be gently homogenized by slowly turning the microtubes upside down and back several times to adequately mix the reagents and, to a large extent, preserve the intactness of the DNA molecules.

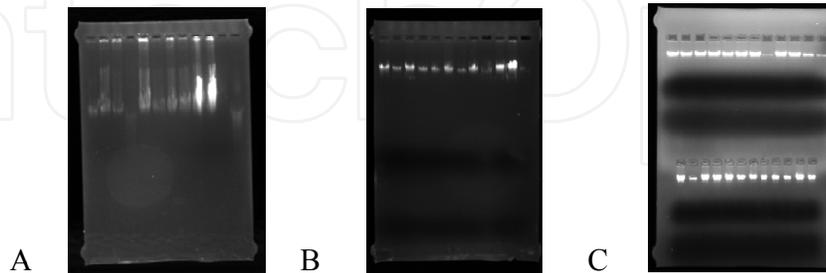


Figure 1. DNA from blood samples. Smeared DNA bands (A), DNA with double-strand breaks (B) and normal intact DNA bands (C). First Report and Final Report of a research grant awarded to Dr. Ivo Sirakov, 2007–2008 – Medicine and Biotechnology – Aids and infectious diseases, Transmissible Spongiform Encephalopathies, funded by the World Federation of Scientists, Geneva, Switzerland.

The methods for preparation of nucleic acid samples can be grouped into thermal extraction, solution-based methods (homemade and commercial kits), column-based methods, and ones that use magnetic particles.

Thermal extraction is a quick and low-cost method that does not require special reagents. It can be used to extract DNA from pure bacterial cultures. There are different variations of the method [8, 9], but they generally include the following procedure: cultures grown 18–24 h are used; dilutions are prepared in Ultra-pure 18.2 MΩ DNase/RNase-free water (when using culture broth, 1:40, or when using agar cultures, 0.5 McF in 1 mL); the samples are heated at 100°C for 5–15 min and then centrifuged; the supernatant is taken and stored at –20°C.

Despite its advantages, this method has some limitations: it cannot eliminate low molecular weight peptides and gives a low 260/280 nm ratio (purity), i.e., about 0.600. This limits the use of the extracted DNA only to conventional PCR.

Thermal extraction is part of the official VTEC *E. coli* diagnostic procedure, which is based on detection of *eae* gene fragments up to 384 bp in length (15). This method, however, appears inapplicable to amplification of larger fragments, especially the *E. coli* 16s RNA gene, which is 1,465 bp in size [10].

The next group of methods, the solution-based ones, in principle, includes the following basic steps: lysis, RNase treatment (if applicable), protein precipitation (two fractions are formed), separation of the fraction that contains the nucleic acids; their precipitation, washing, drying, and regeneration.

One of the earliest and most common solution-based methods is phenol–chloroform extraction. It also has various modifications. Volkin and Carter [11] first developed a method for

RNA extraction with 2 M guanidine hydrochloride, chlorophorm, and alcohol, in which guanidine acts as a deproteination agent and protects the RNA molecules by denaturing the proteins and RNases [12]. After a series of different steps, there is an extraction step using guanidinium thiocyanate–phenol–chloroform [13]. The phenol–chloroform combination (in a 1:1 ratio) gives better protein denaturation (by forming two fractions following centrifugation: a bottom organic phase and an upper aqueous phase) and reduces the amount of poly(A)+ mRNA in the organic phase as well as of insoluble RNA–protein complexes in an intermediate phase [14]. What is more, chloroform prevents the retention of water in the aqueous phase (water can degrade RNA molecules; see above), which results in higher yield [15]. To avoid foam formation, isoamyl alcohol can be added (chlorophorm–isoamyl alcohol, 24:1). It is the acidic properties of phenol that actually determine the partitioning of DNA and RNA in a separate phase: at neutral and slightly alkaline pH (i.e., pH 7–8), DNA and RNA remain in the aqueous phase, since the phosphate diesters are negatively charged. At lower pH (optimal pH 4.8), DNA partitions in the bottom phase, whereas RNA remains in the aqueous phase. This is due to the fact that the phosphate groups in DNA are more prone to neutralization compared to those in RNA [16, 17]. This method is commonly used for DNA extraction from liquid samples, although it also gives good results with tissue cultures, cell cultures etc., provided that they are first homogenized and disintegrated using a lysis buffer (commercial kit or homemade, for example – 10 mM Tris, 1 mM EDTA (ethylenediaminetetraacetic acid) and 0.1 M NaCl) + 20–50 µL of 10–20 mg/mL proteinase K and 1–18 h of incubation at 50–60°C.

There are some specifics in RNA extraction procedures: to protect RNA molecules against RNase attack, chaotropic salts are added to extraction buffers. High-purity RNA can be obtained by guanidine treatment followed by gradient ultracentrifugation in CsCl [12], cesium trifluoroacetate, or LiCl. A disadvantage is the need to use an ultracentrifuge and the fact that the method is time-consuming (it takes about 16 h).

In the case of notoriously difficult samples, Birnboim [18] recommend the use of a combination of SDS (sodium dodecyl sulfate) and urea to more effectively inhibit leukocyte RNases.

During extraction, beta-mercaptoethanol can be added to denature RNases by reducing disulfide bonds and to aid the release of RNA from RNA–protein complexes.

For RNA extraction, phenol and guanidinium isothiocyanate mixtures (Trizol reagent) are available under different commercial brand names and chlorophorm is then added in the extraction process to separate the phases [19].

As a whole, solution-based methods, and especially their homemade versions, have the disadvantage that they use toxic reagents and, therefore, a fume hood is a must; moreover, precision in phase separation is difficult to achieve; the extracted nucleic acid samples are of lower purity (especially in homemade methods) compared to those obtained by column-based methods; the obtained samples have limited use – mainly for conventional PCR. Commercial kits and protocols, albeit validated for downstream applications, do not solve the problem of reagent toxicity and subjective factors related to the person who performs the procedure.

The main advantage of these methods is that they allow for a proportionally larger initial material to be used.

Column-based methods generally use the fact that DNA and RNA molecules are negatively charged to capture them by silica membrane and ion exchange chromatography. These methods include lysis in one or two steps (in silica membranes – chaotropes and proteolytic lysis, and in anion exchange – detergent and enzymatic digestion); then, loading the liquid sample onto a column and centrifugation for the purpose of nucleic acid binding in a high-salt environment; treatment of the column with lysis buffer and/or directly with wash buffer (containing alcohol and high salt); free centrifugation to eliminate the remaining buffer (ethanol*); elution with ultra-pure RNase/DNase-free H₂O (≈pH 8.0) or low-ionic-strength buffer in the case of silica membrane and higher-salt buffer (not suitable for most downstream applications) in the case of anion exchange; incubation of room temperature for 1–3 min and centrifugation. It is possible to repeat the washing and elution steps to achieve further purification (if the column is treated with DNase or RNase after the first elution) or to elute a higher amount of nucleic acids. The standard procedure takes about 20 min per sample or 35–40 min in the case of DNase or RNase treatment. It may be possible to use more starting sample than recommended but it should be kept in mind that this could overload the purification column, resulting in lower yield and/or higher percentage of impurities.* Ethanol remaining in the end-product may lead to the escape of the nucleic acid from the wells during gel electrophoresis and may block restriction analysis, PCR, and other enzyme reactions.

To directly extract mRNA, without an initial procedure of total RNA extraction, oligo(dT) affinity chromatography is used. In this method, oligo(dT) is bound to cellulose or paramagnetic particles. The essence of the method lies in binding of mRNA poly(A⁺) tails to the oligo(dT) fragments attached to the matrix in high-salt conditions. Then, high-salt wash buffer is applied, followed by low-salt wash buffer, and elution under low ionic strength. This method, however, is only limited to eukaryotic cells, since their mRNA is polyadenylated. Another drawback is that the samples may be selectively enriched in mRNAs with shorter poly(A) tracts. What should be avoided is using a maximum amount of sample and overloading, as indicated by higher viscosity and a mucus-like look of the medium.

Another approach to nucleic acid extraction is based on magnetic particles, i.e., the so-called charge-based method: at pH ≤ 6.5, the surface of magnetic particle is positively charged and binds nucleic acids. Depending on the purpose, the unwanted nucleic acid can be eliminated by DNase or RNase treatment. Then, the nucleic acid molecules remaining bound to the particles are released by changing the pH: at pH ≥ 8.5, the surface is neutral and the nucleic acid molecules are released.

In the case of DNA extraction, the comparative analysis of all these methods shows that the quantity/quality of extracted DNA is inversely proportional to the extraction time.

It is commonly accepted to assess the purity of extracted nucleic acid samples based on the ratio of nucleic acids to proteins (impurities). This ratio is determined by measuring the absorbance of the samples at 260 and 280 nm and calculating the 260/280 nm ratio (the

absorbance at 260 nm reflects the mean absorbance of purines and pyrimidines [20, 21]. In the case of DNA samples, it is recommended to make a background correction at 320 nm, which accounts for turbidity. The sample purity reflects the degree to which different contaminants have been eliminated in the nucleic acid extraction procedure. There are different possible sources of contamination. These include the reagents – salts and residual buffer (especially the alcohol-containing wash buffer), as well as different compounds present in the starting material – polysaccharides, phenolic compounds, and DNA or RNA and proteins – nucleases. Elimination of these contaminants is an intrinsic part of the procedures in all the nucleic acid extraction methods. Moreover, there may be included additional purification steps: ultracentrifugation to eliminate high molecular weight polysaccharides; use of beta-mercaptoethanol, dithiothreitol, sulfite, etc.; RNase or DNase can be added and care should be taken to protect the samples from exogenous RNases or DNases of different nature (the samples themselves, human skin, or the laboratory environment), by following the principles of Good Laboratory Practice. It is also recommended to use DEPC, which inactivates RNases in solutions, with the exception of solutions that contain primary amines, such as Hepes buffer and Tris, as they reduce its effect. It is noteworthy that autoclaving will not destroy RNase activity. In plant samples, contamination is often due to polysaccharides and polyphenols, which are eliminated by using polyvinylpyrrolidone. DNA and RNA are also considered contaminants: in DNA analysis, RNA acts as a contaminant and vice versa. Hence, DNase or RNase is used to eliminate contaminating DNA or RNA, respectively. This is so, first of all, because both DNA and RNA contribute to the total nucleic acid content, which – if too high – may block downstream PCR. On the other hand, in gene expression analysis, the methods are sensitive to DNA contamination. For example, reverse transcription (RT) and microarray methods require high RNA purity, since small DNA fragments may anneal to the primers, giving a false positive result, whereas other contaminants (phenol, ethanol, and salts in RT) may react with enzymes, blocking the reaction or increasing the background signal. Conversely, methods such as Northern blotting are not as sensitive to contamination. Thus, sample purity may vary in a certain range – from 0.600 to over 3.0 (mostly for RNA) – depending on the aim of analysis, the extraction method used, the starting material, and the operator. Samples are considered to be of good quality if their 260/280 nm ratio is over 1.7, which is satisfactory for most downstream applications.

However, it is not only the quality of the extracted nucleic acids that is considered important but also their quantity. Depending on the downstream application, if the concentration of nucleic acid in a sample is low, this may, at least in part, be compensated for by using a greater volume of sample in the reaction mixture (the sample purity should be considered as well). For example, if there is insufficient concentration of nucleic acid, non-specific products may be amplified in PCR; or some expected fragments may appear missing in restriction enzyme analysis; or short read lengths may be generated in sequencing. Nucleic acid concentration that is too high may also have adverse effects: amplification of non-specific fragments in PCR or lack of product due to reaction blocking; retention of nucleic acid in gel wells during electrophoresis; incomplete digestion in restriction enzyme analysis (which may be compen-

sated for by adding a proportionally higher amount of enzyme and/or extending the incubation time); or high background in sequencing procedures.

To concentrate, purify DNA and reduce the salt content in DNA samples, precipitation is used (sample: 99% molecular grade ethanol (or isopropanol) 1:1 + 2 to 5% 3.5–7 M ammonium acetate) with 30–60 min incubation at -20°C , centrifugation at 4°C at maximum speed, washing in maximum volume of 70% ice-cold ethanol (-20°C), centrifugation, drying for 3–7 min and resuspending in a desired volume of ultra-pure water, $\frac{1}{4}$ TE buffer or $1\times$ TE buffer, depending on the downstream application and the expected duration of storage.

Another way to determine the quality and quantity of extracted nucleic acids, apart from spectrophotometric analysis, is by gel electrophoresis (GE), which is informative of fragmentation and presence of impurities. Although, in some cases, GE may be sufficient when the researcher is experienced, it is still recommended to use both methods together.

The type of storage and its duration are crucial for the downstream applications. In the case of DNA samples, both the temperature storage and the buffer composition are important factors. Storage at -20°C in $1\times$ PCR buffer for 100 days gives very good results [22]. There are reports that DNA stability can be enhanced by adding 50% glycerol, which limits the formation of ice crystals [23]. Overall, storage at -20°C in commercially available elution buffers (EB) gives stable DNA for use even after a year of storage (EDTA as a component of EB protects DNA against degradation), provided that repeated freezing and thawing are avoided. RNA samples are stored at -80°C in stabilizing buffers that contain EDTA; even storage microcapsules have been developed [1].

3. Nucleic acid and Restriction Enzyme Analysis (REA)

Restriction analysis is an easy-to-perform, inexpensive, and relatively fast method for the study of point mutations and identification of methylated regions in DNA. It can also be used for restriction profiling of micro- and macroorganisms and can serve as a basis for phylogenetic analysis.

In the case of fragmented nucleic acids, e.g., viruses with a segmented genome (Rotavirus), direct fingerprinting is applied, in which individual nucleic acid segments – due to different mobility in an electric field – are distributed at a different distance in an agarose or polyacrylamide gel [24].

Direct electrophoresis, however, does not work in the case of non-segmented nucleic acids. That is why restriction enzymes (restrictases) are used. Restriction enzymes cut the nucleic acid molecule at a specific nucleotide sequence that they recognize. The method is both applicable to total homogeneous (from a single species) DNA and to PCR amplification products. The requirements for the quality of the nucleic acid sample are laid out above (see nucleic acid extraction methods). Additionally, it is recommended to purify the PCR amplifi-

cation product in a gel (most kits recommend 2% gel, although 1.5% gels give better purification) or directly by a column, before restriction. Thus, if there is not enough DNA product in the reaction mixture, more DNA template sample can be added instead of water. It is essential to keep the enzyme/buffer/reaction volume ratio specified in the instructions provided by the enzyme manufacturer.

4. Nucleic acid detection

Detection of nucleic acids and/or traces of them is widely applied in various areas (e.g., biodiversity assessment, marker-assisted selection, molecular diagnostics of infectious diseases, and genetic disorders, etc.), as well as in a range of other fields of industrial and social importance, e.g., food and pharmaceutical industry, healthcare, forensics, etc., only to name a few. For the purpose of nucleic acid detection, there have been developed a number of methods based on hybridization (such as in situ hybridization, molecular beacon) and polymerase chain reactions (PCR, reverse transcription-PCR, real-time PCR).

Nucleic acid hybridization is based on the ability of two complementary nucleic acid strands, at specific conditions, to form a stable double helix. This is mediated by purine–pyrimidine base pairing through hydrogen bonds as first described by Watson and Crick [25]. When hybridization is employed for experimental purposes, a synthetic nucleic acid fragment, the so-called probe, is prepared such that it is labeled (tagged) with a molecule that is easy to detect (the so-called reporter). Reporter molecules were initially radioisotopes, until, in 1981, Langer et al. [26] introduced non-isotopic labeling methods using avidin–biotin binding (covalently bound to the C-5 position of the pyrimidine ring), fluorescent or chemiluminescent dyes. There are now commercial ready-to-use probes for specific diagnostic purposes.

What marked a real turning point in molecular biology was the development of polymerase chain reaction (PCR) by Saiki et al. [27], which basically includes direct in vitro synthesis and multiplication (amplification) of a specific target DNA sequence enclosed between two synthetic oligonucleotides.

Various modifications and versions of PCR have been developed since. This part of the chapter discusses the basic principles underlying conventional PCR, real-time PCR, and reverse-transcription PCR. The mechanisms of different PCRs are illustrated in Figure 2 (Since the structure of nucleic acids is described in greater detail in other chapters, it is only roughly sketched here to show the underlying principles of the reactions).

In order to design an efficient and cost-effective PCR procedure, it is essential to properly choose the reaction components and their precise concentrations: Taq DNA polymerase, buffers, deoxynucleoside triphosphates (dNTPs), MgCl₂, DNA template, and oligonucleotide primers [28, 29]. It is the primers [30] and Taq DNA polymerase [31] that are considered the most important factors that determine the sensitivity and effectiveness of the protocol. Another

key factor is the manufacturer, accounting for different formulations, assay conditions, and/or unit definitions [32].

A standard PCR mixture should include the following main components:

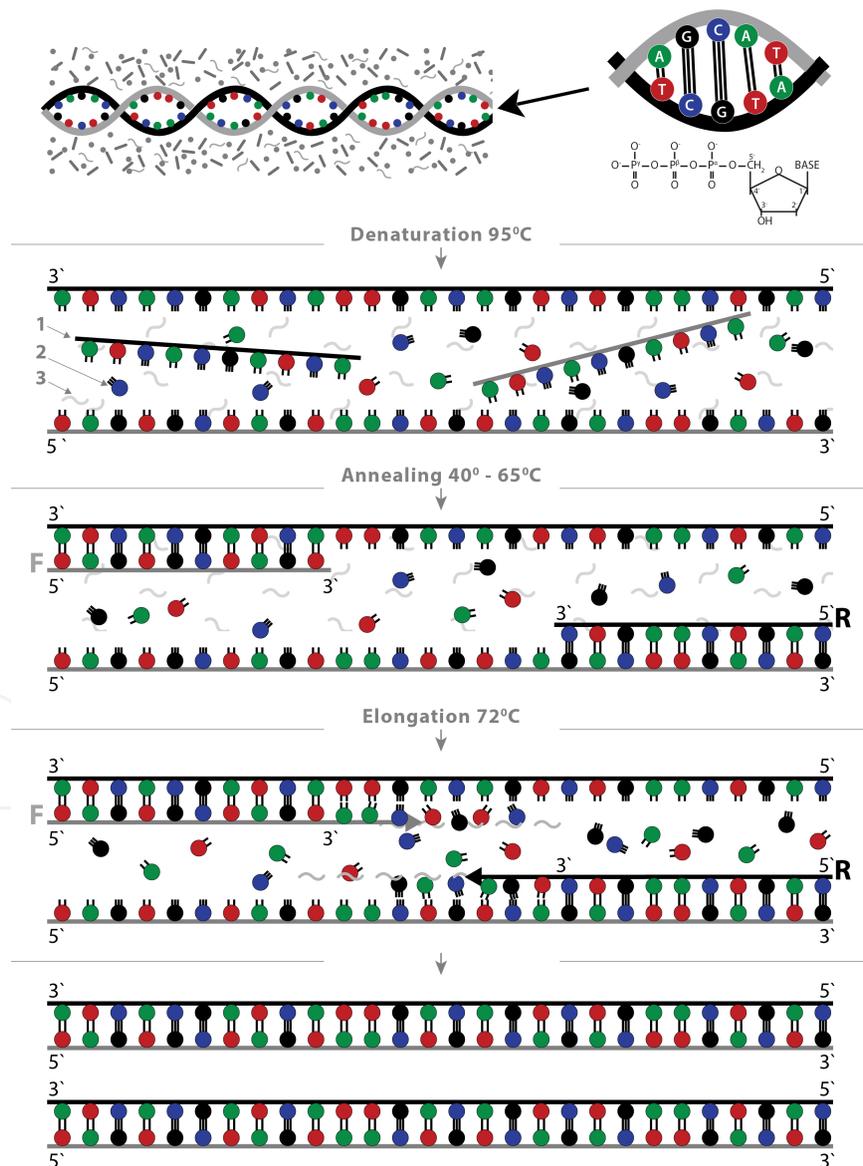
Template DNA – its quality requirements are described in the “**Nucleic acid extraction methods**” part; *Primers* – specific or random complementary oligonucleotides of a different length: a forward primer for the 3′–5′ DNA strand and a reverse primer for the 5′–3′ strand. The primers are particularly important for the reaction sensitivity [30]. That is why, if you are planning to use primers reported by other authors, it is essential to first check their sequences for complementarity and completeness. (It is more often than not that erroneous primer sequences may be published, even in some prestigious journals.) Another point to consider when designing the primer sequences is the GC content, which should be about 50%; the two primers should also have similar melting temperature (T_m) and should not be complementary to one another but only to the target sequence, which should be conservative. In some cases, there may be differences in the sequences targeted by the primers due to mutations (in the genomes of viruses and bacteria). It is then recommended to consider all possible combinations of primer sequences, using the nucleotide coding system for mixed bases, e.g., K (G or T), Y (C or T), etc. **Such differences in the sequences targeted by the primers are used for detection of single-nucleotide polymorphisms (SNP) by real-time PCR** (see below). The primer stability depends on the degree of complementarity and the type of bonds at both ends: for example, the 3′ end should be unstable (to aid the polymerase activity), and the 5′ end should be stable [33], i.e., should contain G or C in the last three bases at the 5′ end. That is why, in some cases, a single-base mismatch in the 3′ end of the primer may not be a problem.

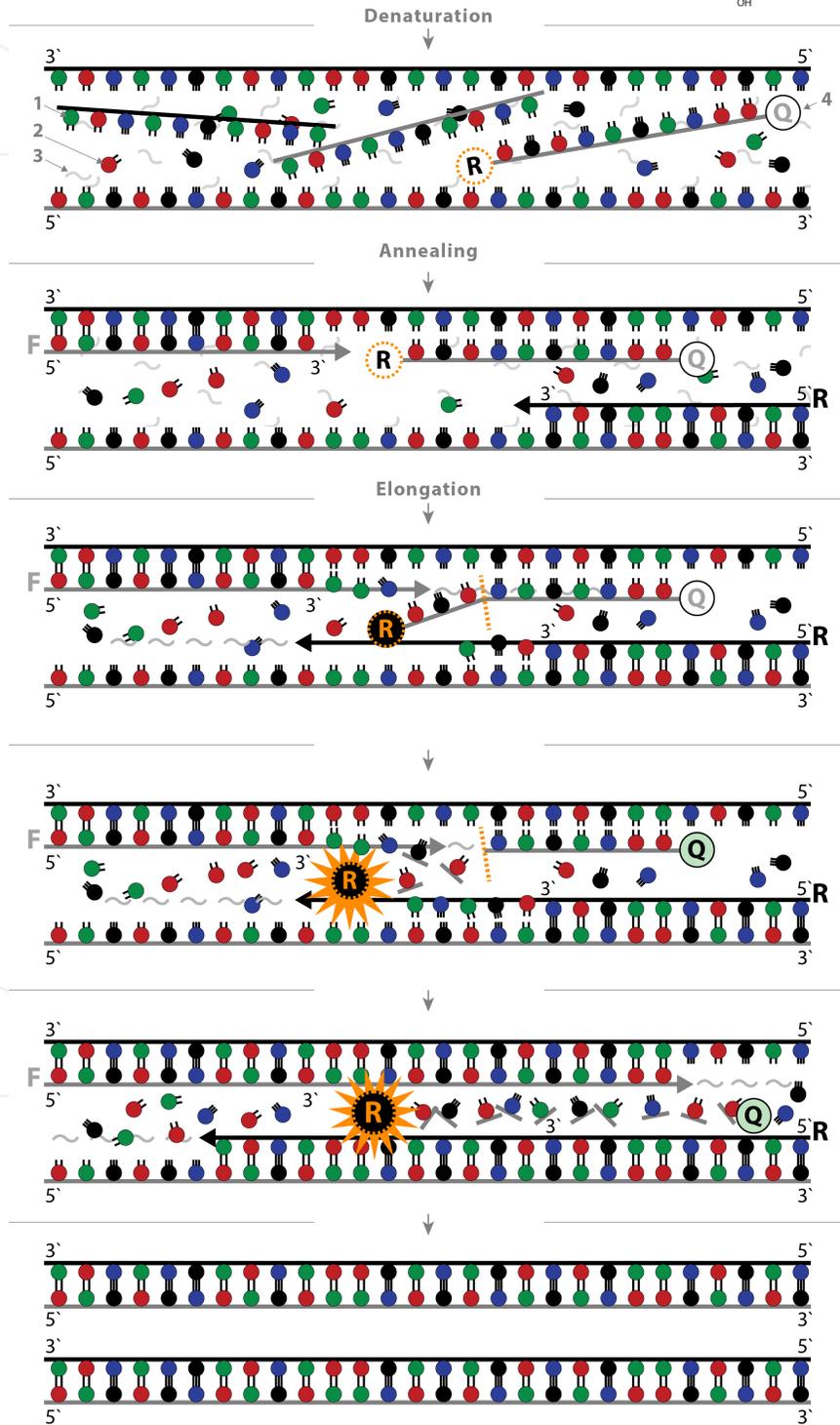
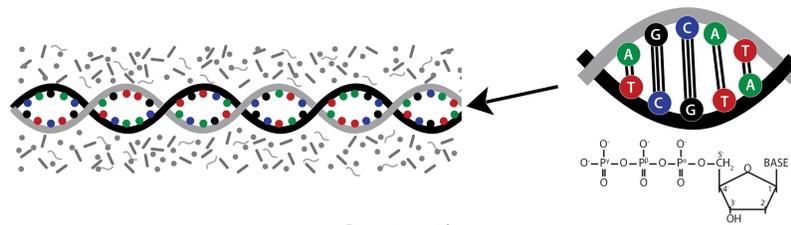
Random primers (RP) – These are short, synthetic, single-stranded DNA segments that are 6 (hexamers) to 10 (decamers) nucleotides in length. They consist of every possible combination of bases. In other words, in the case of hexamer primers, there must be $4^6 = 4,096$ different combinations. Because of that, RP can anneal to any section of the nucleic acid template. The RP approach was described in the late 20th century and is both applicable to analysis of RNA [34, 35] and DNA [36]. The technique based on RP later evolved into RAPD–PCR (random amplified polymorphic DNA), which is a powerful typing method for bacterial species and is also commonly used in construction of genetic maps and fingerprinting libraries and identification of molecular markers [37–39]. (For details see the cited references.)

Taq polymerase is a DNA polymerase from the bacterium *Thermus aquaticus*. It has served as a basis for development of different polymerase enzymes: long range, which allows for incorporation of nucleotides up to 5–10 kb; high fidelity, which includes proofreading exonuclease activity capable of repairing mismatches introduced during strand elongation. The choice of polymerase depends on the method and downstream applications: multiplex PCR, colony PCR, low-copy PCR assay, for difficult (GC-rich) templates, cloning, library preparation, genotyping, etc.

Buffer system (Tris-HCl, $(\text{NH}_4)_2\text{SO}_4$, K/NaCl, MgSO_4) including *deoxynucleoside triphosphates* (dNTPs – dATPs, dGTPs dCTP, and dTTPs), MgCl_2 and 18.2 MΩ DNase/RNase-free H_2O – It

serves as the PCR medium. $MgCl_2$, and particularly Mg^{++} , plays a role in the elongation step as a polymerase cofactor. Additionally, Mg^{++} , along with other cations present in the mixture, reacts with the negatively charged dNTPs (four oxygen atoms surrounding the phosphorus atom (Fig 2) and DNA [40–41]. High salt concentration will lead to non-complementary annealing of DNA strands or to an increase in the DNA denaturation temperature. The buffer also plays a role in maintaining a stable pH in the reaction mixture. PCR products are identified by gel electrophoresis in 1× TBE or 1×TAE buffer and 1.5–3.0% agarose gel. To visualize the results, ethidium bromide is added to the gel (at a concentration of 1 $\mu g/mL$). Ethidium bromide binds DNA non-specifically, which allows the DNA fragments to be visualized by UV illumination. Other dyes that non-specifically bind to DNA, e.g., SYBR Green and others, can also be used.





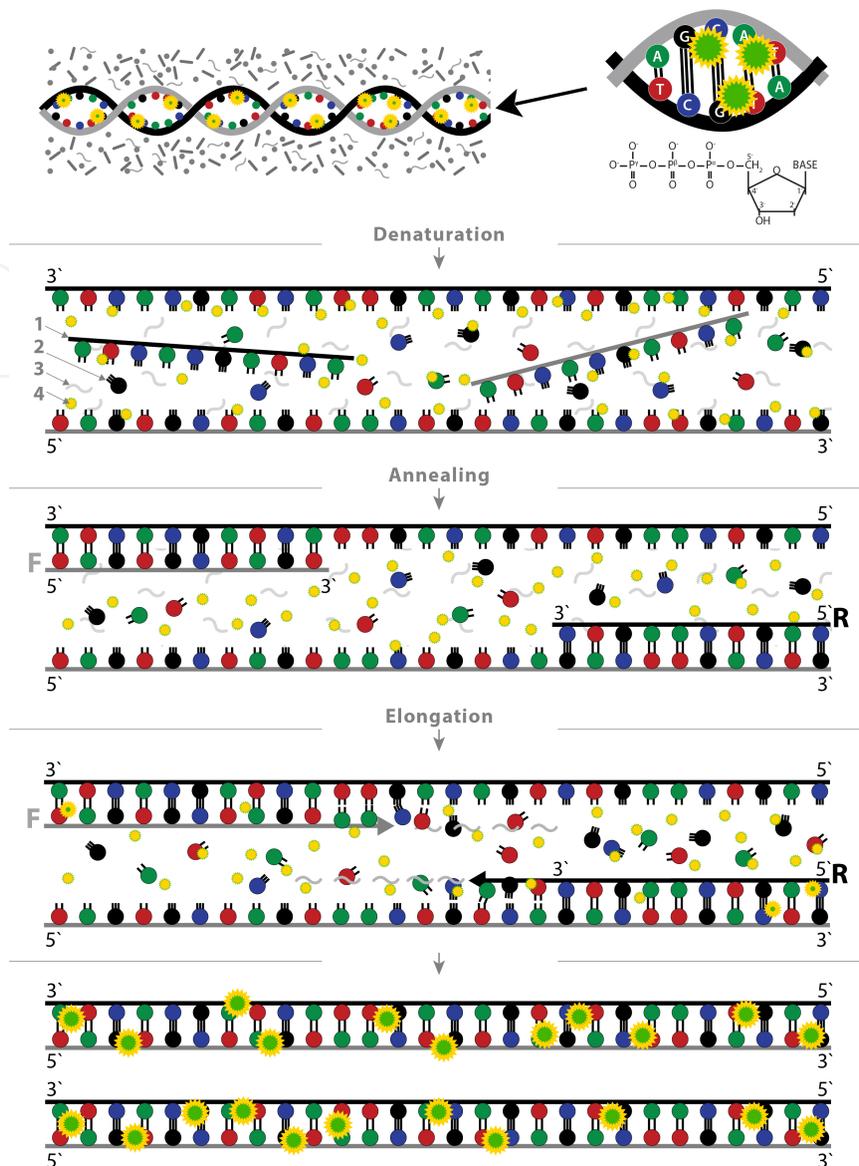


Figure 2. Polymerase chains reactions (PCRs). Components of reactions: 1 – Primers forward and reverse; 2 – deoxynucleoside triphosphates - dATPs, dGTPs dCTP, and dTTPs; 3 – Taq polymerase; 4 – probe labeled with reporter (R) and quencher (Q) molecules; 5 – SYBR Green dye; (A) Conventional PCR includes the following steps: an elongation cycle at 95°C for 1–5 min to activate the polymerase → DNA or cDNA denatures (melts) at 95°C → complementary sequences with specific melting temperature anneal to each other at 40–65°C for 30–120 s → nucleotides are incorporated in the growing strand and the target sequence is amplified at 72°C for up to 60 s. After 35 cycles, a single dsDNA copy is amplified into 2^{36} copies; (B) TaqMan real-time PCR (real-time PCR Taq has 5′–3′ exonuclease activity) 3(b) underlying principle of TaqMan real-time PCR. The reaction mixture includes the same main components as conventional PCR plus a synthetic oligonucleotide (probe) that is labeled with a reporter and a quencher and is complementary to an internal region in the 3′–5′ strand of the sequence of interest (4). The reaction can also include two steps per cycle – denaturation at 95°C and annealing/elongation at 60°C [8, 45]; and (C) Underlying principle of SYBR Green real-time PCR. SYBR Green dye (5) non-specifically binds to dsDNA. During denaturation, the dye is released; it then binds again to the PCR fragment during the elongation step. SYBR Green emits fluorescence when bound to dsDNA. Thus, the more fragments are amplified, the stronger the fluorescence intensity will be. The signal is graphically recorded the same way as in the TaqMan reaction.

Particular attention should be paid to some ambiguities that may arise from the usage of similar acronyms to denote different PCR techniques: the acronym RT is only used to denote reverse transcription; real-time PCR is not abbreviated, and quantitative real-time PCR is commonly denoted as real-time qPCR, whereas real-time reverse-transcription PCR is typically denoted as real-time RT-PCR, and in quantitative analysis, as real-time RT-qPCR.

Fragment amplification in real-time PCR is based on the same principle as conventional PCR and includes the same basic steps. The difference lies in the method of detection, which needs specially designed equipment. Real-time PCR is based on detection of the fluorescence emitted by a reporter molecule in real time, which is associated with another synthetic oligonucleotide (probe) that is complementary to an internal sequence of the target gene and is labeled with a reporter (R) and a quencher (Q) molecule. The signal emitted by the R molecule is detected after the probe becomes detached from the complementary strand and the R molecule is released by hydrolysis (Figure 2B) – TaqMan version [42]. A signal is emitted and detected in the so-called LightCycler version – by increase and detection of fluorescence resonance energy transfer, via hybridization of R and Q side by side [43–44]. These detection approaches laid the foundations for development of the so-called quantitative real-time PCR (qPCR), which is widely used in infectious disease diagnostics (e.g., human hepatitis viruses), SNP genotyping and allelic discrimination, somatic mutation analysis, copy number detection/variation analysis, chromatin IP quantification, DNA methylation detection, RNA analyses – gene and miRNA expression studies. In this case, the signal is monitored in the course of amplification (i.e., during the early and exponential accumulation of the PCR product) to detect the first significant peak in the amount of PCR product, which is proportional to the initial quantity of target template.

Real-time PCR results are visualized as curves on a graph that reflects the accumulation of signal (Figure 3). The result is obtained based on a pre-prepared standard curve and an internal, positive and negative control that need to be run; i.e., gel electrophoresis is not needed, but may be used as an exception, in case of equipment malfunction, to detect the products.

SYBR Green real-time PCR – This fluorescent dye was first used for detection of nucleic acids in agarose gels [46] and was later introduced in real-time PCR amplifications [44, 46, 47]. The method is based on the fact that the dye **only** binds to double-stranded DNA, which is accompanied with an increase in fluorescence. Thus, the signal intensity correlates with the amount of amplified DNA fragment and, respectively, with the initial sample input amounts (Figure 2C). In 2004, Hubert et al. [48] described the SYBR Green molecule as [2-[N-(3-dimethylaminopropyl)-N-propylamino]-4-[2,3-dihydro-3-methyl-(benzo-1,3-thiazol-2-yl)-methylidene]-1-phenyl-quinolinium]. SYBR Green-based analysis can be used in amplification of any dsDNA and does not require a probe, which makes it less costly. However, the SYBR Green dye may yield false positive signals, as it intercalates into any dsDNA, including non-specific dsDNA sequences.

In general, real-time PCR is more sensitive than conventional PCR and needs the target sequences to be shorter than those used in conventional PCR, maximum 300–400 bp in length; results are obtained in real time and it is not necessary to use gel electrophoresis. The cost of a single reaction (excluding the controls) is much higher than that of conventional PCR. For

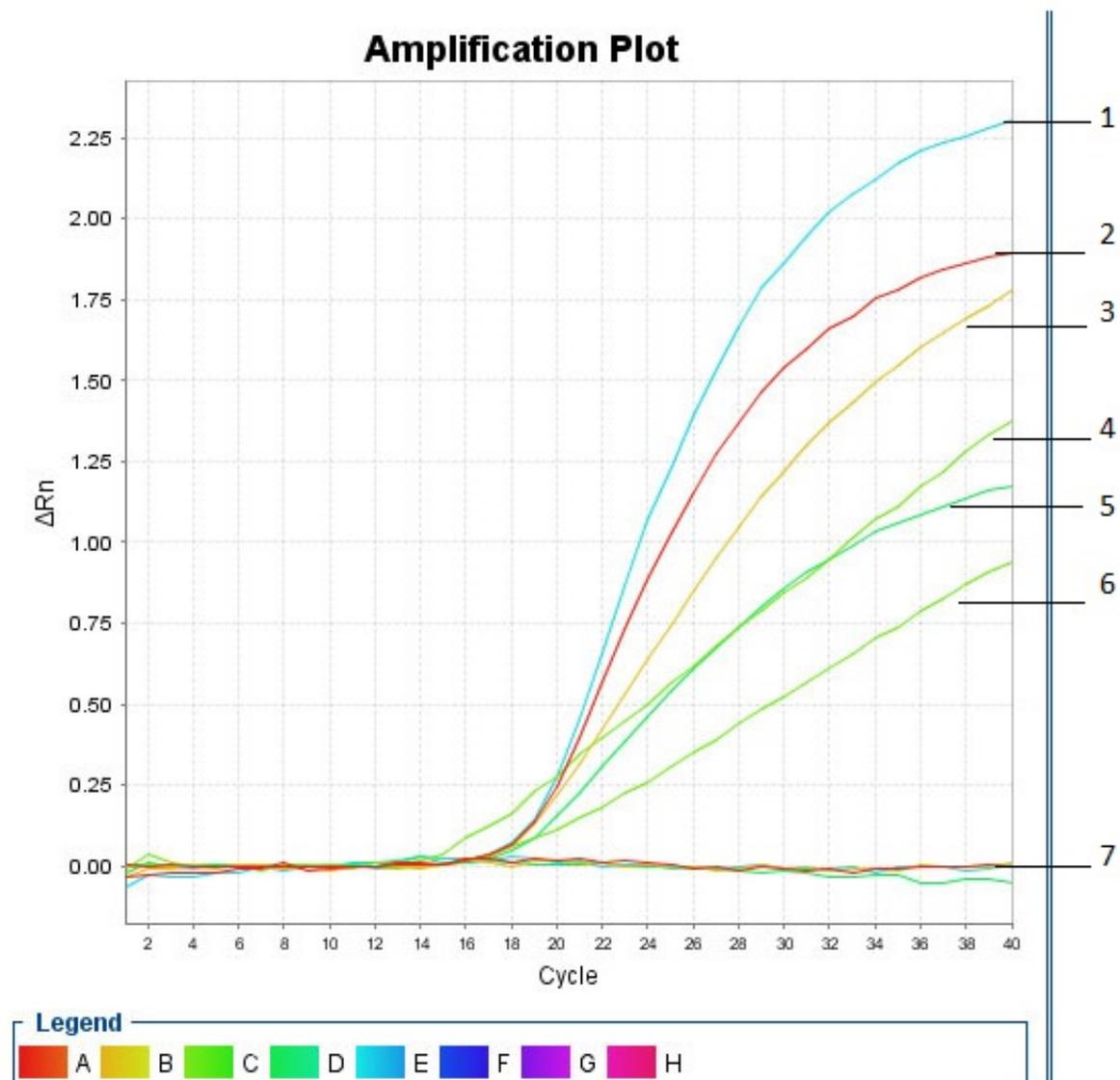


Figure 3. TaqMan real-time PCR data for detection of Shiga toxin genes in DNA extracted from *E. coli* culture broth [45]. 1–6 – amplification curves of internal controls; 7 – lines indicating non-amplification of negative controls and two samples.

example, in conventional PCR, the minimum reaction cost is 0.70 euro, and in TaqMan real-time PCR, about 3 euros (including tips, tubes, and gloves).

Reverse-transcription (RT) PCR is specific in that it includes an additional reverse transcription reaction generating cDNA. This cDNA is then used in conventional or real-time PCR, either in one step (the reaction directly proceeds from reverse transcription to subsequent amplification steps in the same tube) or in two steps (the RT reaction is run separately and a new reaction mixture is prepared for conventional or real-time PCR).

These methods are also applied for multiplex reactions, i.e., amplification of different target sequences in one and the same reaction. What is important for the primer pairs used in the reaction is for them not to be complementary to each other so that they do not form dimers. Another key point is for the primer pairs to have similar annealing temperature (Conventional PCR can tolerate 1°C difference in the annealing temperature of each primer.).

Other methods that are also based on amplification of a target nucleic acid sequence are ligase chain reaction (LCR), nucleic acid sequence-based amplification (NASBA), and strand displacement amplification (SDA).

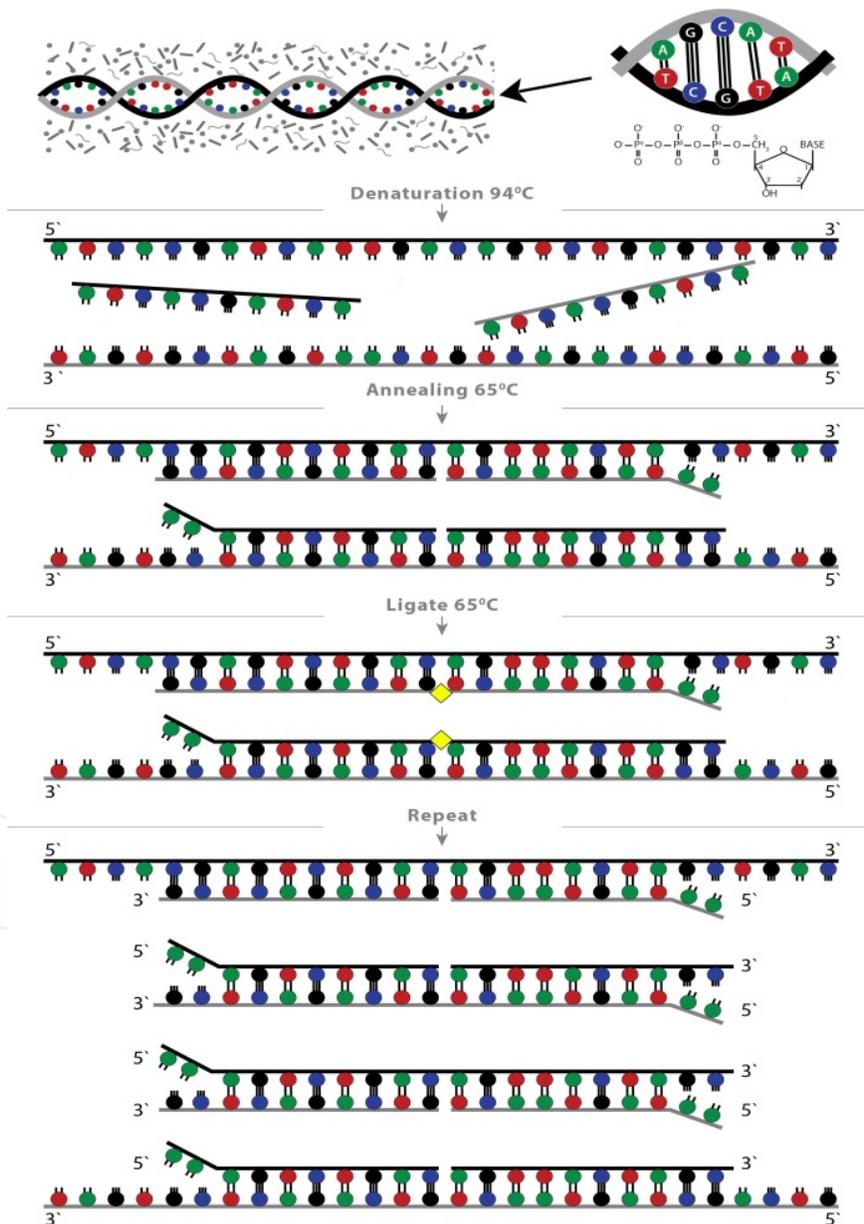


Figure 4. Ligase chain reaction (LCR).

Ligase chain reaction (LCR) was first described by Barany [49]. It combines a ligase reaction with amplification and is particularly suitable for differentiation of single-base substitutions (Figure 4). LCR is based on the following principle: DNA is denatured at 94–95°C and four primers are annealed to the complementary strands at ~65°C, i.e., ca. 5°C below their T_m . The thermostable ligase, then, proceeds to ligate only those primers that share perfect complementarity to the target sequence and hybridize immediately next to each other. Thus, if two primers bear a single base-pair mismatch at the junction, they will not ligate effectively enough and, in turn, there will be no product amplification. To avoid ligation of the 3' ends, the discriminating primers contain a 2-bp non-complementary AA tail at their 5' ends [49, 50].

The nucleic acid sequence-based amplification (NASBA) method is based on isothermal amplification used for RNA detection. Briefly, its principle and reaction mixture are as follows: Primer 1, which is complementary to the 3' end of the target RNA (+) strand and includes a T7 promoter sequence, anneals to it. Next, reverse transcription yields a cDNA (-) strand. Then, the hybrid RNA-cDNA strand is separated via destruction of the RNA strand by RNase H. In the next step, primer 2 anneals to the 3' end of the DNA (-) strand and reverse transcription yields dsDNA that contains a T7 promoter. Then, T7 RNA polymerase generates RNA (-) copies, primer 2 anneals to the 3' end of the RNA (-) strand, RT yields copyDNA (+) and primer 1 anneals to copyDNA (+), followed by RT and dsDNA synthesis, which can serve as a template for RNA (+) or RNA (-) synthesis to close the cycle [51].

Strand displacement amplification (SDA) combines the principles of isothermal DNA amplification with those of restriction enzyme digestion [52]. The reaction contains the following components: four primers, DNA polymerase, REase HincII, dGTP, dCTP, dTTP, dATP α S, and takes about 2 h. Basically, in SDA, the primer has two parts: the 5' end includes a specific HincII enzyme cleavage site (-G-T-T-G-A-C-) and the 3' end is complementary to the target DNA sequence. DNA polymerase generates a complementary strand and a thiophosphate modification is incorporated – deoxyadenosine 5'-[α -thio]triphosphate (dATP[α S]), in the specific HincII enzyme cleavage site (-C-A_s-A_s-C-T-G-). Then, HincII cleaves the strand at its specific site (-G-T-T ↓ G-A-C-), but not at the complementary sequence (-C-A_s-A_s-C-T-G-). The free 3'-OH group serves as a starting point for strand elongation by DNA polymerase. Thus, the specific HincII cleavage site is regenerated, which closes the cycle.

There are various modifications and versions of these methods, even some commercial kits; however, it is not possible for all of them to be discussed here.

5. Nucleic acid identification by sequencing methods

Sequencing is the basic method to determine the nucleotide sequence of DNA and RNA molecules. There are different sequencing methods and variations, as the methodology evolves based on successful adaptation and application of the properties of nucleic acids.

Sanger method – in 1974, Sanger et al. [53] reported a sequencing method now known as the Sanger method. It is similar with fragment analysis in that the reaction includes a DNA

template, radioactive labeled dNTPs and dideoxynucleotides (ddNTPs – without 3'-hydroxyl group, which is essential in phosphodiester bond formation – ddA, ddG, ddC, and ddT), T7 DNA polymerase (with ability to incorporate 2',3'-dideoxynucleotides), a primer (forward or reverse) and reaction buffer. The annealing, labeling, and termination steps are performed on different thermoblocks, and the polymerase reaction at 37°C. The polymerase enzyme can incorporate either dNTPs or ddNTPs (depending on their relative concentration) at each elongation step. Elongation will proceed if dNTPs are added and will stop if a ddNTP is added at the 3' end of the strand. The resulting fragments are of different size (length) and, in gel electrophoresis, will migrate toward the positive electrode at a rate of migration inversely proportional to their molecular weight. The method can differentiate fragments that are only 1 bp different in length.

This served as the basis for development of another approach, in which fluorescent dyes are used instead of radioactive isotopes – cycle sequencing or capillary sequencing (the term “capillary” stems from the fact that electrophoresis is performed in a special matrix in capillary tubes, and fluorescence is detected by means of a laser beam). The reaction components are the same as those in the Sanger method, but are mixed in such a way as to allow thermal cycling: denaturation, annealing, elongation, and generation of a balanced population of short and long fragments, using the same principle as the Sanger method (Figure 5A). Specialized software is used to process the detected fluorescence of each fragment and to plot the result as an electropherogram (Figure 5B) in ABI, FASTA, and PAUP file format.

A limitation of the method is the size of the fragments that can be sequenced: maximum 800–1000 nucleotides per run.

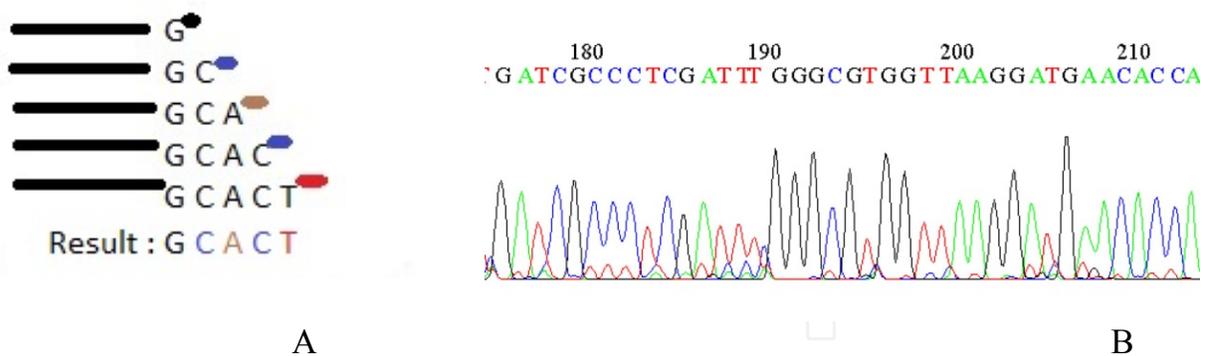


Figure 5. Generation of fragments by incorporation of dNTPs and ddNTPs (labeled) in the Sanger method and cycling sequencing (A). Sample electropherogram (portion) of the beta-lactamase OXA-48 gene of *Klebsiella pneumoniae* strain OXA48BG, NCBI, GenBank accession number KJ959619.1 [54] (B).

The sequences obtained using forward and reverse primers are analyzed by programs available either as freeware [55, 56] or as commercial software. What is important to remember is to always check the sequencing results in the file generated by the software against the electropherogram: there may often be discrepancies between the nucleotide sequence in the

file and that in the electropherogram. In such cases, the electropherogram should be considered more reliable but the background effect should also be accounted for.

In the alignment of sequences (first, between the F and R primers of a sample and, second, between different samples), it is important for them to be equal in length. The next step is sequence analysis – phylogenetic analysis, genetic distance analysis, etc. For example, in phylogenetic analysis, it is particularly important to choose the mathematical model that is most appropriate for each particular case. Instead, there are software programs especially designed to determine the most appropriate model depending on the sequence length, the number of sequences, potential substitutions, etc. For example, jModelTest [57] analyzes and selects among 89 different models.

5.1. Next-Generation Sequencing (NGS) platforms

The way is now open for nucleic acid research using NGS technologies. This part of the chapter will focus on three NGS platforms: Illumina, Nanopore, and Ion Torrent.

5.2. Illumina sequencing by synthesis

The Illumina technology is based on the principle described above, i.e., incorporation of fluorescently labeled dNTPs by DNA polymerase using a DNA or a cDNA template in a sequence of cycles. The identification of incorporated nucleotides is based on fluorophore excitation. The whole process includes the following steps:

- Nucleic acid preparation is one of the critical steps (it may vary depending on the operator and the consumables/method used). It is important for the target DNA (total DNA, PCR products, cDNA) to be of the highest possible purity.
- Sequencing library preparation (another critical step) is done by random fragmentation of the target DNA, followed by 5' and 3' ligation of fragments by adapters. Thus, each strand in the DNA double helix is adapter-ligated. Fragmentation and ligation may be performed in a single step, which, according to (9), markedly increases the effectiveness. The adapter-ligated fragments are then amplified by PCR and purified in a gel.
- Cluster generation takes place on the surface of a flow cell. The flow cell is, in fact, a glass slide with channels on whose surface two types of oligonucleotides are attached. These oligonucleotides are complementary to one of the ends of the library fragments (the adapter) and are identical with the other end. This ensures cluster generation is done by hybridization of the adapter-ligated amplified fragments from the library with synthetic oligonucleotides on the surface of the flow cell that are complementary to the adapters. Each single-stranded fragment binds with the complementary end, while its other end remains free, perpendicular to the surface of the flow cell (The free end is identical with the oligonucleotide in the flow cell to provide synthesis of a complementary sequence in the new strand.). Then complementary strand synthesis can begin, displacing the original strand when the synthesis is completed. The fragments thus bound act as templates for the so-called bridge amplification of discrete clonal clusters. In bridge amplification, the free end of a new single-

stranded DNA fragment anneals (by hybridization) to a complementary oligonucleotide on the flow-cell surface (resulting in an \cap -shaped bridge-like structure); then, a DNA polymerase enzyme synthesizes a complementary DNA strand by incorporation of unlabeled nucleotides. The resulting double-stranded DNA template is denatured, which releases the two single-stranded DNA fragments both from one another as well as from the flow-cell surface at one end: if the initial ssDNA fragment had its 5' end free, then it is restored and the new copy remains bound at its 3' end (which is complementary to the free 5' end). Thus, the templates generated by bridge amplification form clusters of single-stranded fragments bound at one end and the number of DNA templates grows. (Then, the reverse templates are removed and only the forward strands remain – clones identical to the original templates)

- Sequencing – The samples are ready for the sequencing step when the clusters are complete. A proprietary reversible terminator-based method is used: a sequencing primer is annealed (it is complementary to the part of the adapter that is adjacent to the DNA fragment of interest), followed by cycles of incorporation of terminator-bound dNTPs. In each cycle, there is a mixture of the four different dNTPs and they naturally compete for incorporation based on complementarity. As a result, only one complementary dNTP per cycle is incorporated in the growing strand. In each cycle, a labeled terminator-bound dNTP is incorporated in each chain in each cluster and is detected based on emission wavelength and intensity. The signal emitted by each nucleotide incorporated in each strand at the end of a cycle is recorded in real time in each cluster. The cycle is repeated n times to obtain an n -base long sequence. When the new strand is synthesized and its sequence is recorded, it is removed. This approach of base-by-base sequencing achieves very high precision, as errors are eliminated even in repetitive sequence regions and homopolymers.
- Data processing and analysis – The fluorescence emissions from different clusters are digitally processed in parallel and are visualized as nucleotide sequences for each individual cluster. Following sequence alignment, the data are compared against referent sequences to perform phylogenetic analysis, single-nucleotide polymorphism analysis, distance determination, and a range of other analyses.

This technique allows multiplex analysis to be carried out by creating distinct libraries based on the so-called index sequences (short nucleotide sequences specific for each library that can be used like a barcode), which are attached during library preparation step. Different libraries are first individually prepared and are then combined together and loaded in one and the same flow cell lane. The labeled libraries are sequenced simultaneously, in a single run of the equipment, and at the end of the process, the sequences are exported in a single file. Next, a demultiplexing algorithm is used to separate the sequences in different files based on their barcode. This is followed by alignment with referent sequences of interest.

The platform is compatible with different library preparation methods depending on the purpose of the sequencing analysis (whole genome sequencing, target sequencing, mRNA sequencing, 16s RNA gene sequencing, etc.). This is possible because the sequencing steps that come after the library preparation step are fundamental and do not depend on the library preparation method. There are ready-to-use, standardized library preparation kits designed

for different sequencing purposes and more and more new approaches and modifications are being developed [58].

5.3. Nanopore technology

Nanopore technology is an innovative NGS platform that also includes sequence library preparation. The library preparation step requires maximal purity of the target nucleic acid (DNA, PCR products, cDNA). This is achieved by magnetic particle purification (It is probably possible to use other methods as well, e.g., column silica-based ones, if the manufacturers are reliable and guarantee high purification of the nucleic acid samples). The target nucleic acid samples are ligated between the 5' end of one of the strands and the 3' end of the other strand of the double-stranded DNA (in the shape of a hairpin). Bayley sequencing technology is used. In this method, the DNA molecule is basically passed through a system consisting of a processive exonuclease enzyme bound to a protein nanopore (Figure 6A). The enzyme unzips the DNA double helix and pushes/cleaves one of the strands through the aperture of the protein nanopore in a base-by-base manner (The aperture is only a few nanometers in diameter.). This continues until the hairpin loop at the end of the first strand is reached and then the enzyme simply proceeds on to push the ligated reverse strand through the nanopore as well. The system is designed so that a constant ion current flows through the aperture of the free protein nanopore. This current is specifically disrupted when each base enters the aperture. These specific disruptions in the ion current are recorded by an electronic device and are interpreted to identify each base. The disruption that each of the four bases causes in the ion current is different due to differences in their chemical structure and chemical characteristics. Recordings from multiple channels in parallel allows high-throughput sequencing of DNA.

Protein nanopores are inserted into an electrically resistant polymer membrane so that a membrane potential can be created driving the ion current through the pore aperture. This ion current is disrupted in a specific way when a nucleotide passes through the aperture of the protein nanopore, which causes a change in the membrane potential.

An array of microscaffolds holds the membrane in which the nanopores are embedded, giving stability to the structure during operation. Each microscaffold on the sensor array chip contains an individual electrode, allowing for multiple nanopore experiments to be performed in parallel. Each nanopore channel is controlled and measured by an individual channel on a corresponding, bespoke application specific integrated circuit (ASIC) (Figure 6B) [59].

Another NGS platform is the Ion Torrent sequencing technology. In this method, following library preparation, the sequencing step is actually performed in wells that contain the DNA template, an underlying sensor, and electronics. It works as follows: when a new nucleotide is incorporated in the growing DNA strand, a proton (H^+) is released, which causes a change in the pH in the well. This leads to changes in the surface potential of a metal-oxide sensing layer and to changes in the potential of the source terminal of the underlying field-effect transistor. This is the signal that a complementary nucleotide has been incorporated at the end of the growing DNA strand. To determine the DNA sequence, the equipment needs to be able to differentiate between the four bases. This is done by flushing each well with one type of nucleotide at a time. If the base is complementary, the nucleotide will be incorporated and a

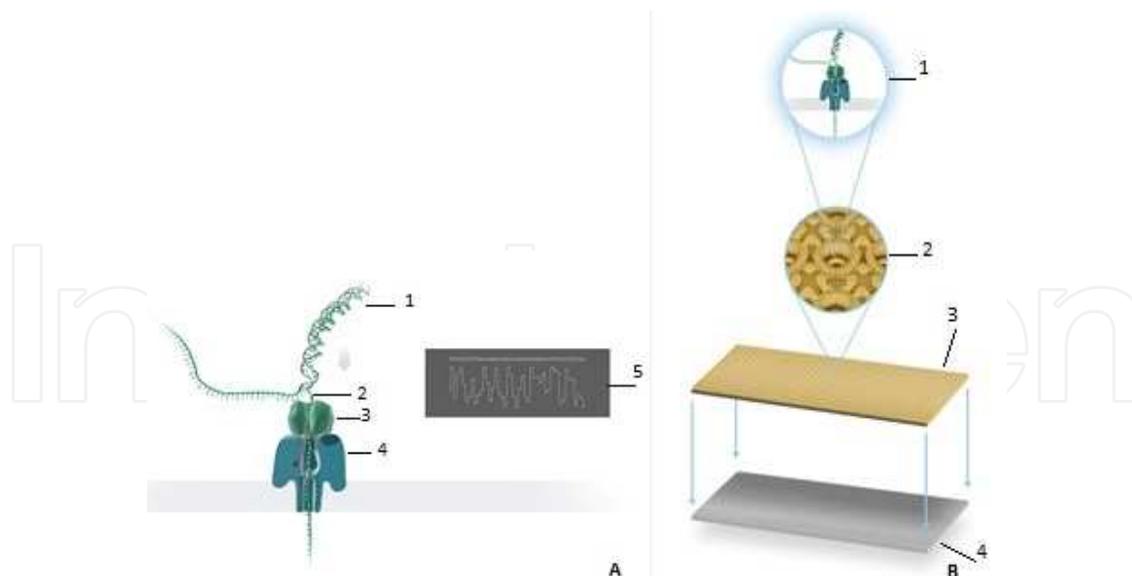


Figure 6. (A) Underlying principle used in nanopore technology sequencing (illustrative data). 1 – DNA library is prepared; 2 – the dsDNA strands are separated and one of the strands passes through 3 – a processive exonuclease enzyme, and 4 – a protein nanopore, in a base-by-base manner, causing a characteristic disruption in the electrical current flowing through the nanopore, which is in fact the signal that is detected in real time. When the end of the first strand is reached, the 3'- to 5'-end hairpin ligation of the two DNA strands allows the processive enzyme to simply continue to push the second strand through the nanopore; 5 – Representative image of DNA sequencing data. (B) 1 – A nanopore protein; 2 – Array of Microscaffold; 3 – Array Chip; 4 – Application Specific Integrated Circuit.

change in the pH will be recorded; and if the base is not complementary, there will be no pH change [60].

NGS technologies are a new trend and yet more approaches and applications in nucleic acid analysis are being developed.

Author details

Ivo Nikolaev Sirakov*

Address all correspondence to: insirakov@gmail.com

Medical University – Sofia, Department of Microbiology, Sofia, Bulgaria

References

- [1] Fabre AL, Colotte M, Luis A, Tuffet S, Bonnet J. An efficient method for long-term room temperature storage of RNA. *Euro J Human Genetics* 2014;22:379–85; doi: 10.1038/ejhg.2013.145

- [2] Poulsen HE, Specht E, Broedbaek K, Henriksen T, Ellervik C, Mandrup-Poulsen T, Tonnesen M, Nielsen PE, Andersen HU, Weimann A. RNA modifications by oxidation: a novel disease mechanism? *Free Radic Biol Med* 2012;52:1353–61.
- [3] Bruskov VI, Malakhova LV, Masalimov ZK, Chernikov AV. Heat-induced formation of reactive oxygen species and 8-oxoguanine, a biomarker of damage to DNA. *Nucleic Acids Res* 2002;30:1354–63.
- [4] Evans RK, Xu Z, Bohannon KE, Wang B, Bruner MW, Volkin DB. Evaluation of degradation pathways for plasmid DNA in pharmaceutical formulations via accelerated stability studies. *J Pharm Sci* 2000;89:76–87.
- [5] Cataldo F. Ozone degradation of ribonucleic acid (RNA). *Polym Degrad Stabil* 2005;89:274–81.
- [6] Seyhan AA, Burke JM. Mg²⁺-independent hairpin ribozyme catalysis in hydrated RNA films. *RNA* 2000;6:189–98.
- [7] Butler E, Gelbart T, Kuhl W. Interference of heparin with the polymerase chain reaction. *BioTechniques* 1990;9:166.
- [8] Identification of the subtypes of Verocytotoxin encoding genes (vtx) of *Escherichia coli* by conventional PCR - EU Reference Laboratory for E. Coli, Istituto Superiore di Sanità, EU-RL VTEC_Method_006_Rev 1; 18.06.2013.
- [9] Sirakov I, Popova R, Daskalov H, Slavcheva I, Gyurova E, Mitov B. Shigatoxin-producing *Escherichia coli* in raw cow milk from small farm producers and phylogenetic subtype determination. *JFAE* 2014;12(3,4):108–14.
- [10] Popova R [PhD thesis] Research on spread, characteristics and pathogenic factors of *Escherichia coli* isolated from food. Sofia, Bulgaria: National Diagnostic and Research Veterinary Institute, Bulgarian Food Safety Agency; 2014.
- [11] Volkin E, Carter CE. The preparation and properties of mammalian ribonucleic acids. *J Am Chem Soc* 1951;73(4):1516–9.
- [12] Chirgwin JM, Przybyla AE, MacDonald RJ, Rutter WJ. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 1979;18(24):5294–9.
- [13] Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 1987;162:156–9.
- [14] Perry RP, La Torre J, Kelley DE, Greenberg JR. On the lability of poly(A) sequences during extraction of messenger RNA from polyribosomes. *Biochim Biophys Acta* 1972;262(2):220–6.
- [15] Palmiter RD. Magnesium precipitation of ribonucleoprotein complexes. Expedient techniques for the isolation of undegraded polysomes and messenger ribonucleic acid. *Biochemistry* 1974;13:3606–15.

- [16] Brawerman G, Mendecki J, Lee SY. A procedure for the isolation of mammalian messenger ribonucleic acid. *Biochemistry* 1972;11(4):1972.
- [17] Puissant C, Houdebine LM. An improvement of the single-step method of RNA isolation by acid guanidinium thiocyanate – phenol – chloroform extraction. *BioTechniques* 1990;8:148–9.
- [18] Birnboim HC1. Extraction of high molecular weight RNA and DNA from cultured mammalian cells. *Methods Enzymol* 1992;216:154–60.
- [19] Rio DC, Ares M Jr, Hannon GJ, Nilsen TW. Purification of RNA using TRIzol (TRI reagent). *Cold Spring Harb Protoc* 2010 Jun;2010(6):pdb.prot5439. doi: 10.1101/pdb.prot5439
- [20] Wyatt GR. The purine and pyrimidine composition of deoxypentose nucleic acids. *Biochem J* 1951;48(5):584–90.
- [21] Clark LB, Tinoco I Jr. Correlations in the ultraviolet spectra of the purine and pyrimidine bases. *J Am Chem Soc* 1965;87(1):11–5. doi: 10.1021/ja01079a003
- [22] Barbara R, Karin F, Claus V, Martin W, Peter R. Impact of long-term storage on stability of standard DNA for nucleic acid-based methods. *J Clin Microbiol Nov* 2010;4260–2. doi:10.1128/JCM.01230-10
- [23] Bonner G, Klibanov AM. Structural stability of DNA in nonaqueous solvents. *Biotechnol Bioeng* 2000;68:339–44.
- [24] Haralambiev H. *Animal Viruses. Diagnostic of virus infection.* Publisher – Pandora, Haskovo, 2002, pp. 58–59.
- [25] Watson JD, Crick FHC. A structure for deoxyribose nucleic acid. *Nature* 1953;171:737–8.
- [26] Langer PR, Waldrop AA, Ward DC. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 1988;239:487. doi: 10.1126/science.2448875
- [27] Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn, GT, Mullis KB, Erlich H. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 1988;239(4839):487–91.
- [28] Linz U, Delling U, Rubsamem W. Systematic studies on parameters influencing the performance of the polymerase chain reaction. *J Clin Chem Clin Biochem* 1990;28:5–13.
- [29] Sarkar G, Kapelner S, Sommer SS. Formamide can dramatically improve the specificity of PCR. *Nucleic Acids Res* 1990;18:7465–8.
- [30] He Q, Marjamäki M, Soini H, Mertsola J, Viljanen MK. Primers are decisive for sensitivity of PCR. *Biotechniques* 1994;17(1):82, 84, 86–7.

- [31] Drummond R, Gelfand DH. Isolation, characterization and expression in *Escherichia coli* of the DNA polymerase gene from *Thermus aquaticus*. *J Biol Chem* 1989;264:6427–36.
- [32] Vinay KS, Anil K. PCR primer design. *Molecul Biol Today* 2001;2(2):27–32.
- [33] Sheffield VC, Cox DR, Lerman LS, Myers RM. Attachment of a 40 base pair G+C rich sequence (GC-clamp) to genomic DNA fragments by polymerase chain reaction results in improved detection of single base changes. *Proc Natl Acad Sci* 1989;86:232–6.
- [34] Taylor JM, Ilmensee IL, Summers J. Efficient transcription of RNA into DNA by avian sarcoma virus polymerase. *Biochim Biophys Acta* 1976;442:324–30.
- [35] Noonan KE, Roninson IB. mRNA phenotyping by enzymatic amplification of randomly primed cDNA. *Nucl Acids Res* 1988;16(21):10366.
- [36] Feinberg AP, Vogelstein B. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* 1983;132(1):6–13.
- [37] Jones C, Kortenkamp A. RAPD library fingerprinting of bacterial and human DNA: applications in mutation detection. *Teratog Carcinog Mutagen* 2000;20(2):49–63.
- [38] Rossetti L, Giraffa G. Rapid identification of dairy lactic acid bacteria by M13-generated, RAPD-PCR fingerprint databases. *J Microbiol Methods* 2005;63(2):135–44.
- [39] Kumar NS, Gurusubramanian G. Random amplified polymorphic DNA (RAPD) markers and its applications. *Sci Vis* 2014;11(3):116–24.
- [40] Cowan JA. Structural and catalytic chemistry of magnesium-dependent enzymes. *Biometals* 2002;15(3):225–35.
- [41] Sreedhara A, Cowan JA. Structural and catalytic roles for divalent magnesium in nucleic acid biochemistry. *Biometals* 2002;September 15,3:211–23.
- [42] Livak KJ, Flood SJA, Marmaro J. Method for detecting nucleic acid amplification using self-quenching fluorescence probe. Applied Biosystems Division, Perkin-Elmer Corp. US Patent US5538848, 1996.
- [43] Cardullo A, Agrawal S, Flores C, Zamecnik PC, Wolf DE. Biochemistry detection of nucleic acid hybridization by nonradiative fluorescence resonance energy transfer. *Proc Natl Acad Sci USA* 1988; Dec 85:8790–4.
- [44] 44 multisample fluorimeter with rapid temperature control. *Biotechniques* 1997;22:176–81.
- [45] Popova R, Taseva R, Sirakov I, Daskalov H. Fecal contamination of natural water reservoirs in Pirin Mountain with *Escherichia coli*. *Tradition and Modernity in Veterinary Medicine*. Issn 1313-4337, University Of Forestry, Sofia, Bulgaria, pp. 98–102, 12/2012.
- [46] Schneeberger C, Speiser P, Kury F, Zeillinger R. Quantitative detection of reverse transcriptase-PCR products by means of a novel and sensitive DNA stain. *PCR Methods Appl* 1995;4:234–8.

- [47] Bengtsson M, Karlsson HJ, Westman G, Kubista M. A new minor groove binding asymmetric cyanine reporter dye for real-time PCR. *Nucleic Acids Res* 2003;31:45.
- [48] Giglio S, Monis PT, Saint CP. Demonstration of preferential binding of SYBR Green I to specific DNA fragments in real-time multiplex PCR. *Nucleic Acids Res*. 2003;31:136.
- [49] Zipper H, Brunner H, Bernhagen J, Vitzthum F. Investigations on DNA intercalation and surface binding by SYBR Green I, its structure determination and methodological implications, *Nucl Acids Res* 2004;32:12, 103. doi:10.1093/nar/gnh101
- [50] Barany F. The ligase chain reaction in a PCR world. *Genome Res* 1991;1:5–16.
- [51] Wiedmann M, Wilson WI, Czajka J, Luo J, Barany F, Batt CA. Ligase chain reaction (LCR)--overview and applications. *Genome Res* 1994;3:S51–S64.
- [52] Guatelli JC, Whitfield KM, Kwoh DY, Barringer KJ, Richman DD, Gingeras TR. Isothermal, in vitro amplification of nucleic acids by a multienzyme reaction modeled after retroviral replication. *Proc Natl Acad Sci USA* 1990;87:1874–8.
- [53] Walker GT, Little MC, Nadeau JG, Shank DD. Isothermal in vitro amplification of DNA by a restriction enzyme/DNA polymerase system. *Proc Natl Acad Sci USA* 1992;89:392–6.
- [54] Sanger F, Donelson JE, Coulson AR, Kösse H, Fischer D. Determination of a nucleotide sequence in bacteriophage f1 DNA by primed synthesis with DNA polymerase. *J Mol Biol* 1974;90(2):315–33.
- [55] Marteva-Proevska Y, Strateva T, Sirakov I, Zlatkov B, Markova B. Emergence of *Klebsiella pneumoniae* carrying OXA-48 carbapenemase in Bulgaria. *International Journal of Biological Sciences and Applications* 2015 x(x): xx-xx
- [56] Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994;22:4673–80.
- [57] Edgar RC. Muscle: multiple sequence alignment with high accuracy and multiple sequence alignment with high accuracy. *Nucleic Acids Res* 2004;32(5):1792–7.
- [58] Posada D. J Model Test: phylogenetic model averaging. *Mol Biol Evol* 2008;25:1253–6.
- [59] Illumina Nextera DNA Library Preparation Kits data sheet. 2004. www.illumina.com/documents/products/datasheets/datasheet_nextera_xt_dna_sample_prep.pdf
- [60] <https://nanoporetech.com/applications/dna-nanopore-sequencing>
- [61] <https://www.thermofisher.com/us/en/home/life-science/sequencing/next-generation-sequencing/ion-torrent-next-generation-sequencing-technology.html>