

# METHODS OF MOLECULAR DIAGNOSTICS FOR FISH SPECIES IDENTIFICATION

Tatyana A. Fomina<sup>1,2\*</sup>, Vladimir Yu. Kornienko<sup>1</sup>, Mihail Yu. Minaev<sup>1</sup>

<sup>1</sup> V.M. Gorbatov Federal Research Center for Food Systems of Russian Academy of Sciences, Moscow, Russia

<sup>2</sup> National Centre for Safety of Aquatic Fisheries Products and Aquaculture, Moscow, Russia

**KEY WORDS:**

*DNA analysis, molecular diagnostics, PCR, sequencing, sampling, species identification, adulteration*

**ABSTRACT**

The growth in demand for fish products as a result of globalization of trade caused a risks and threats of selling poor-quality and falsified fish products. This has become a great problem both for supervising agencies and for consumers.

Many countries have regulations on food labelling and safety. For example, in the Russian Federation, Republic of Belarus and Republic of Kazakhstan has been passed the Technical Regulation of the Customs Union TR CU022/2011 “Food products in part of their labeling” that aims to prevent misinformation of consumers to ensuring realization of consumer rights to reliable information about food products, and Technical Regulation TR EAEU040/2016 “On safety of fish and fish products” requires indication of the zoological name of the species of the aquatic biological resource or the object of aquaculture.

Fish species identification is traditionally carried out based on external morphological traits. However, it becomes impossible to identify species by ichthyological traits upon fish cutting, if the head and fins are removed, and the body is cut on pieces (especially, in case of fillets) and even more so upon technological processing. In this case, objective analytical methods of species identification are used, which are based on ELISA or PCR. However, DNA-based methods have several advantages compared to ELISA methods and complement traditional morphological identification methods. This paper gives a wide overview of the most recent and used methods of fish species identification based on DNA analysis such as single-strand conformation polymorphism (SSCP) analysis, species-specific PCR, real-time PCR, polymerase chain reaction-restriction fragment length polymorphism analysis (PCR-RFLP), DNA barcoding, Sanger sequencing and next-generation sequencing (NGS).

**FUNDING:** The article was published as part of the research topic No. 0437.19.001 of the state assignment of the V. M. Gorbatov Federal Research Center for Food Systems of RAS

## 1. Introduction

Fish is one of the most saleable goods in the international food market. Globalization of fish trade and an increase in the demand for fish products have led to serious concerns about the risk of selling poor-quality and even adulterated fish products. These concerns were caused by the revealed high level of incorrect labeling of fish products worldwide, which led to deterioration of their quality and safety [1,2,3,4,5].

The risk and threat of adulteration of edible fish products linked with unreliable labeling of species origin of their composition, especially in technologically processed products, became a serious problem both for controlling agencies and for consumers [6,7].

Legislation regarding food safety assessment and labeling was enacted in many countries of the world [6].

In the EU countries, products are labeled according to Regulation (EU) No 1169/2011 of the European Parliament and of the Council of 25 October 2011 on the provision of food information to consumers [8].

In the Russian Federation, Republic of Belarus and Republic of Kazakhstan, Technical Regulation of the Customs Union TR CU022/2011 “Food products in part of their labeling” is in force to prevent misinformation of consumers in reference to ensuring realization of consumer rights to reliable information about food products [9].

On one hand, this Technical Regulation 022/2011 requires reliable indication of finished product composition, while Technical Regulation TR EAEU040/2016 “On safety of fish and fish products” requires indication of the zoological name of the species of the aquatic biological resource or the object of aquaculture; on the other hand, the presence of allergens should be in-

dicated on a label of any food product even if they are contained at a trace level.

Special concerns that led to strengthening measures on quality control of produced and distributed food products are linked with possible consequences of eating food that contains substances causing allergic reactions or intolerance, when these substances are absent on a label of a consumed product [7].

Over the last years, an increase in the number of patients with food allergy has been observed; this problem is more topical for childhood. At present, fish and fish products are among the most frequent causes of food allergy both in children and in adults, which is linked with the widespread increase in consumption of these products. The prevalence of fish allergy widely varies and is about 0.2% in total population [10].

The majority of allergic reactions on fish is caused by the main allergen — sarcoplasmic protein parvalbumin, which is present in many species of river and sea fish. Up to 90% of “fish” allergy patients react exactly on this protein. Parvalbumin is thermally stable and remains in a product even after its cooking. In addition to parvalbumin, other fish allergens were revealed — tropomyosin, collagen, aldolase, enolase, vitellogenin, calcitonin [11,12].

Fish allergy causes pathological reactions, which are based on immune mechanisms — specific IgE-mediated reactions. Clinical manifestations of allergy can be quite various including rhinitis, angioedema, urticaria, gastrointestinal disorders (nausea, vomiting) and also the most severe, life-threatening form — the anaphylactic shock [13,14,11].

Correct labels of fish-based food products can play a part in stimulation of sustainable fishing helping consumers rightly detect fish origin and, therefore, allowing them to make justified and responsible buying decisions [5].

These circumstances underline the necessity to control correspondence of labeling and reveal cases of species adulteration in fish and fish products.

To detect fish species adulteration, it is necessary to identify fish. Identification of fresh and raw fish can be carried out visually by characteristic taxonomic morphological traits, such as shape and pattern of scales, shape of the body and its size, shape and number of fins, location of eyes, specific features of internal organs and so on. However, it becomes impossible to identify species by ichthyological traits upon fish cutting, if the head and fins are removed, and the body is cut on pieces (especially, in case of fillets) and even more so upon technological processing [15, 16].

In this case, other methods of identification are necessary – instrumental. In the world practice, raw material and product identification is based on species specificity of both protein molecules and DNA. These methods allow avoiding adulteration when expensive species are replaced with cheaper ones [15,17,18].

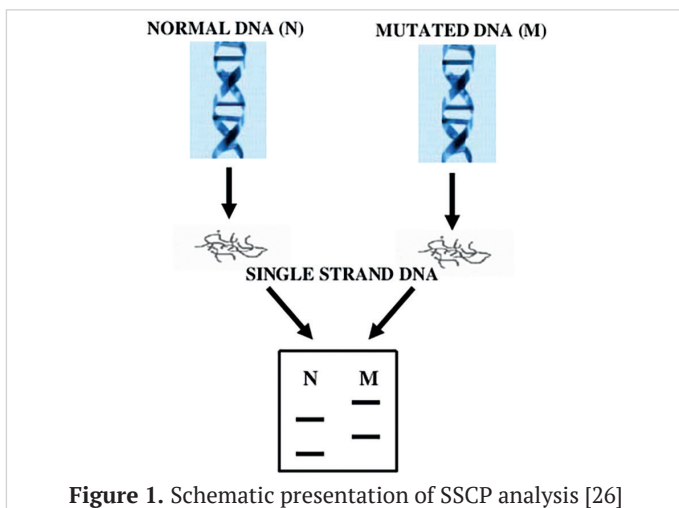
DNA-based methods have several advantages compared to protein analysis. First of all, the DNA molecule is more stable to an impact of high temperatures and even if DNA is partly destroyed due to product technological processing (cunning, cooking), it can still be used for investigations, for example, by PCR, which allows amplification of small DNA fragments with sufficient information for species identification. Secondly, DNA is present in all tissue types and it can be extracted from any organic material. In addition, methods of DNA analysis are preferable due to larger variability of the genetic code. For example, the mitochondrial genome is characteristic by pronounced instability – the evolution rate of mtDNA exceeds that of nDNA by 10–20 times. The evolution rates of the mitochondrial genome of mammals are estimated as  $5.7 \times 10^{-8}$  substitutions per synonymous site per year. This instability ensures intra- and interspecies polymorphism, which allows the most effective use of mtDNA to differentiate closely related species of animals, fish and birds [7,16,19,20,21,22,23].

This paper discusses currently available methods for fish species identification.

**2. Main part**

**2.1. Single-strand conformation polymorphism (SSCP) analysis**

SSCP (single-strand conformation polymorphism) – analysis of conformation polymorphism of single-stranded DNA (ssDNA) – is a method for detection of differences in ssDNA electrophoretic mobility due to mutations by spatial organization (conformation) of molecules. Conformation of small ssDNA depends on the composition of nucleotides; therefore, substitution of even one nucleotide leads to changes in the spatial structure (Figure 1) [24,25].

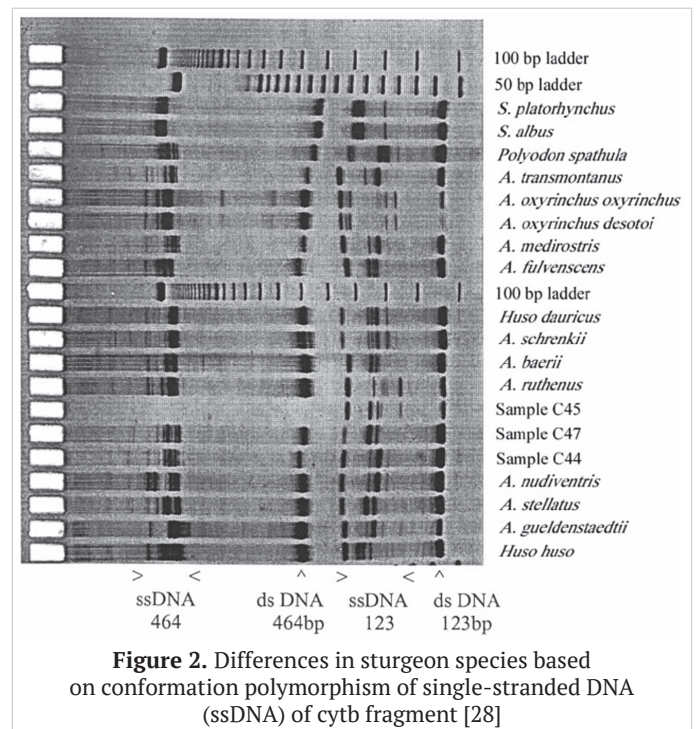


**Figure 1.** Schematic presentation of SSCP analysis [26]

Thus, detection of the changed pattern of ssDNA fragment migration in different samples in the SSCP analysis allows suggesting species differences, even if species are closely related [16].

The application of this method is described in several studies. For example, Weder, J. et al (2001) [27] used the SSCP method, which had been initially applied to identify tuna of genera *Katsuwonus* and *Thunnus*, to study other species of fish and animals. The 148 bp amplicon obtained using PCR of the fragment of cytochrome b gene (cytb) was used for the study. The ssDNA fragments of blue ling, carp, haddock, mackerel, mackerel shark, saithe, catfish, Alaska pollack, and skipjack gave two to four clear patterns; however, they were different from those obtained with ssDNA samples from tuna (*Thunnus*). The ssDNA fragments of other fish species showed weak bands (cod, spined dogfish) or their absence (Atlantic salmon, halibut, herring, pike-perch, plaice, redbfish, sprat, trout).

Rehbein (1997) used the SSCP analysis for identification of different fish species from the family *Acipenseridae*. The length of the amplicon of the cytochrome b gene (cytb) fragment was 358 bp (Figure 2) [28].



**Figure 2.** Differences in sturgeon species based on conformation polymorphism of single-stranded DNA (ssDNA) of cytb fragment [28]

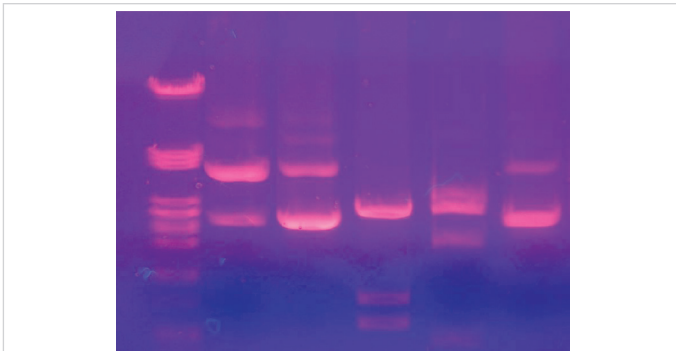
SSCP analysis is rapid and easy to use; nevertheless, this method shows three main disadvantages: a) it is necessary to run a reference sample and test sample simultaneously on the same gel, b) intraspecies variation can lead to different conformations, which, in turn, can lead to incorrect identification, c) sometimes two bands with different intensities can be seen. The reason for this can be the fact that ssDNA exists in several states of conformation depending on electrophoretic conditions [16].

**2.2. Species-specific PCR**

Species-specific PCR is one of the most common variants of using PCR (polymerase chain reaction) method for the diagnostic purpose. The prerequisite in this method is the knowledge of nucleotide sequence of the gene, on which basis the species identification will be carried out; that is, primers will be designed [29,30].

Upon corresponding reaction conditions, such primers generate a fragment that is visualized by agarose gel electrophoresis only in the presence of DNA of this species (Figure 3). This

procedure is applicable only when some previous knowledge about material analyzed is available and identification is to be made [16].



**Figure 3.** Photo of gel containing template DNA (photo from the authors' archive)

This method has been applied for quite a long time. For example, Vadim J. Birstein et al. [31] describe the use of species-specific PCR for identification of black caviar from beluga sturgeon (*H. huso*), sevruga (*A. stellatus*) and Russian sturgeon (*A. gueldenstaedtii*). For each of these species, a set was developed, which consisted of a pair of specific primers complementary to the regions of template DNA, between which the target sequence was located. In these studies, the occurrence of an amplification product obtained using primers, was assumed to enable correct identification of species under investigation [28,31].

More recent studies of the same groups of scientists showed that species-specific PCR did not allow accurate differentiation of Russian sturgeon (*A. gueldenstaedtii*) from closely related species (*A. baerii*, *A. naccarii* and *A. persicus*) due to overlapping mitochondrial DNA profiles [32].

Although this method has its advantages being easy to use, inexpensive and rapid for species identification of beluga sturgeon (*H. huso*) and sevruga (*A. stellatus*), it failed to differentiate between caviar obtained from Russian sturgeon (*A. gueldenstaedtii*) and Persian sturgeon (*A. persicus*) [28].

There is also a possibility to use together several pairs of species-specific primers in a single reaction tube for simultaneous amplification of DNA of different species. This modification was given the name multiplex PCR, which is one of the variants of species-specific PCR [33]. Micheli E. et al. [34] describe the development of one-step analysis based on triplex-polymerase chain reaction (PCR) to discriminate between three tuna species, yellowfin (*Thunnus albacares*), bigeye (*Thunnus obesus*), and skipjack (*Katsuwonus pelamis*), even in highly processed food samples such as canned or cooked tuna. In this analysis, amplification of the specific regions occurs only in the presence of template DNA, and the species origin of the template DNA is assessed by the size of the PCR product: yellowfin has 245 bp, bigeye 262 bp and skipjack 113 bp [34].

Species-specific PCR is based on unique interspecies nucleotide differences; nevertheless, there are aspects that can lead to false positive or false negative results, which require inclusion of reference samples in each analysis [28,16,33].

### 2.3. Real-time PCR

Real-time PCR is a modification of conventional PCR, where accumulation of amplification products is analyzed by a special apparatus, which distinctive feature is a possibility to excite and detect fluorescence that reflects accumulation of amplicons in each amplification cycle [35].

In this variant, the fluorescent signal comes from another important component added to the reaction mixture — the DNA

fragment that contains a fluorescent dye and a fluorescence quencher (a probe) attached at the 5' and 3' ends of the oligonucleotide, respectively. The probe is complementary to one of the chains within the amplicon and in the course of copying by polymerase of the DNA fragment specified by the primer, the probe is degraded due to 5'-3'-exonuclease activity of polymerase. The dye and fluorescence quencher are spatially separated and fluorescence occur. Therefore, occurrence of one amplicon is linked with fluorescence of one molecule of free (not quenched) fluorophore. A probe for the marker gene is made in one color, for example, fluorescein — FAM, a probe for a gene under investigation is made in another color, for example, introducing rhodamine (R6G, ROX and others) into oligonucleotide and so on. Thus, the fluorescent signal over the course of PCR increases proportionally to the quantity of an amplification product and allows observing the process of product accumulation over the course of the reaction using computer software. A moment of pronounced increase in a signal and its separation from the baseline, the so-called threshold cycle (log-phase), depends on the initial quantity of target DNA. The higher the quantity of DNA in a sample, the earlier the beginning of an increase in the fluorescence signal is observed [29,36,37,38].

Real-time PCR, also known as quantitative real-time PCR (qPCR), is used for measuring gene copies or a level of gene expression [29]. Hird et al. describe the study of the variation in the proportion of muscle tissue to the numbers of a single copy gene in haddock (*Melanogrammus aeglefinus*) and, then, the development of a haddock-specific quantitative assay using TaqMan technology and real-time PCR for a single copy gene. The study shows that the calibration curve was able to quantify model samples with the accuracy of up to 7% of the true percentage [39].

The accuracy of PCR quantification to a large extent depends of the reference material used in construction of the standard curve [16].

Sánchez A. et al. [40] describe the use of real-time PCR for absolute, absolute-relative and relative quantification of the most valued hake species in European markets, European hake (*Merluccius merluccius*). The authors compared two systems, nuclear and mitochondrial. The research showed that the best quantification results for this species in binary mixtures with non-target species (*Merluccius capensis*) and using the species-specific real-time PCR system were achieved with a relative quantification approach. It was demonstrated that absolute quantification using the nuclear system was appropriate for the quantification of the *Merluccius genus* in food model samples [40].

Real-time PCR is the most common technology to use for species identification. Continuous measurement of fluorescence allows eliminating stages, which usually are necessary after performing PCR, that is, electrophoresis and gel staining. Moreover, in case of real-time PCR, potential risks of contamination are significantly reduced as samples with the reaction mixture are sealed throughout the analysis [29].

### 2.4. Polymerase chain reaction-restriction fragment length polymorphism analysis (PCR-RFLP)

PCR-RFLP (Polymerase chain reaction-restriction fragment length polymorphism analysis) is a method, where a fragment of the studied gene carried the recognition site for endonuclease is amplified with its following cutting with the corresponding enzyme leading to appearance of several smaller fragments with different sizes (restriction fragments). Restriction fragment sizes are analyzed by gel-electrophoresis [41,42].

PCR-RFLP allows identifying different meat types originated from mammals, birds or fish [43] and has been widely used in many countries worldwide.



Georgina L. Hold et al. [44] describe the study on application of the PCR-RFLP based method for identification of salmon species in food products. A 464 bp portion of the cytochrome B gene was used as a target sequence for amplification; restriction fragments were obtained using the following enzymes: Dde I; Nla III; Hae III; Bsp 1286I; Eco RII; Sau 3AI. The reliability and practicality of the method were tested in the inter-laboratory study, in which five European laboratories took part: 1. Rowett Research Institute, Greenburn Road, Aberdeen, AB21 9SB, Scotland, UK; 2. Bundesforschungsanstalt für Fischerei, Institut für Biochemie und Technologie, Palmaille 9, 22769 Hamburg, Germany; 3. Departamento de Bioquímica y Biología Molecular, Facultad de Biología, Universidad de Santiago de Compostela, 15706 Santiago de Compostela, Spain; 4. Instituto de Investigaciones Marinas (CSIC), Eduardo Cabello 6, 36208 Vigo, Spain; 5. Instituto de Investigacao das Pescas e do Mar, Avenida de Brasilia, Lisbon, Portugal. Ten samples of morphologically identified salmon species (*Salmo salar*, *Oncorhynchus keta*, *Oncorhynchus kisutch*, *Oncorhynchus gorbuscha*, *Oncorhynchus nerka*, *Oncorhynchus tshawytscha*, *Oncorhynchus mykiss*, *Salvelinus alpinus*, *Salvelinus fontinalis*, *Salmo trutta*) and two encoded samples (1 – *S. salar* (commercial product), 2 – a mixture of two salmon species *O. keta* and *O. gorbuscha*) were used as the control samples. All results (including those for encoded samples) showed 100% agreement and were correctly identified. In addition to the inter-laboratory study, larger scale investigations of UK commercial products were also carried out covering the whole range of available salmon products. In almost all cases, the declared salmon species was confirmed; a trout species was detected in one product, which label declared only the presence of salmon. The performed research confirmed the reproducibility of the method in different laboratories as well as its applicability for analysis of commercial products [44,45].

Lin and Hwang [46] used this method to detect eight species of the family Scombridae: bluefin tuna (*T. thynnus*), albacore (*T. alalunga*), bigeye tuna (*T. obesus*), yellowfin tuna (*T. albacares*), skipjack (*Katsuwonus pelamis*), eastern little tuna (*E. affinis*), frigate mackerel (*Auxis thazard*) and oriental bonito (*Sarda orientalis*) in samples of canned tuna. Two sets of primers were designed to amplify 126 bp and 146 bp fragments of mitochondrial cytochrome B gene, and five restriction enzymes (Bsp1286I, HincII, RsaI, ScaI and MboII) were used to analyze short length fragments. The method was successfully applied for authentication of 18 samples of canned tuna [46].

Ya-Jung Wu [47] used the PCR-RFLP method to identify filefish species (*Monacanthidae*), which are a food delicacy in Taiwan. The cytochrome B gene region with the molecular weight of 465 bp was chosen for primers. The obtained results showed that when using this method, six commercial filefish species could be identified not only in fresh but also in thermally processed products [41].

The PCR-RFLP method is regarded as robust and easy to use for identification of fish species; however, problems with analysis robustness due to the use of individually prepared and non-optimized components as well as the manual nature of the analysis can potentially affect the reliability of results [48].

Dooley J. J. [49] and his colleagues presented an optimized PCR-RFLP approach for fish species identification. The scientists replaced the gel-electrophoretic steps for fragment separation, detection and analysis with a chip-based capillary electrophoresis system with the use of the Agilent 2100 Bioanalyser. The presented solution reduces analysis duration and allows obtaining a result with a minimum impact on a sample. Figure 4 presents the scheme of the analysis.

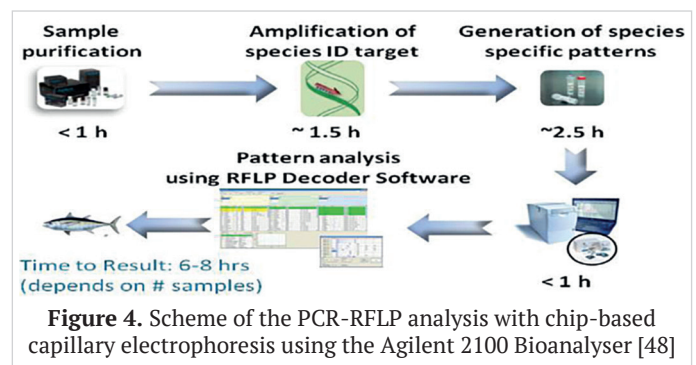


Figure 4. Scheme of the PCR-RFLP analysis with chip-based capillary electrophoresis using the Agilent 2100 Bioanalyser [48]

The main steps include extraction of genomic DNA, amplification of the target DNA — the region of the cytochrome B gene, cleavage of the PCR product with three different restriction enzymes (Dde I; Nla III; Hae III) and separation on the bioanalyser; then, the stage of the instrumental analysis is applied using software with extensive database of experimental profiles of fish species Agilent RFLP Decoder for RFLP pattern analysis.

Software analyzes the obtained result for a test sample and compares it with profiles of authentic fish species in the database using standard calculation methods to reveal the most probable matches. The flexibility of software allows easy addition of users' profiles to extend the number of species that can be identified based on the experimental models. In addition, analysis of mixed products is envisaged, including different fish species as well as detection of the presence of pork, beef, mutton and turkey. At present, the further work on adaptation of the method to specific tasks for fish species identification has been carried out, but even now the reagent kit and consumables for detection of fish species Agilent DNA Fish Species ID Ensemble are available for purchase including the territory of the Russian Federation [49,48,50].

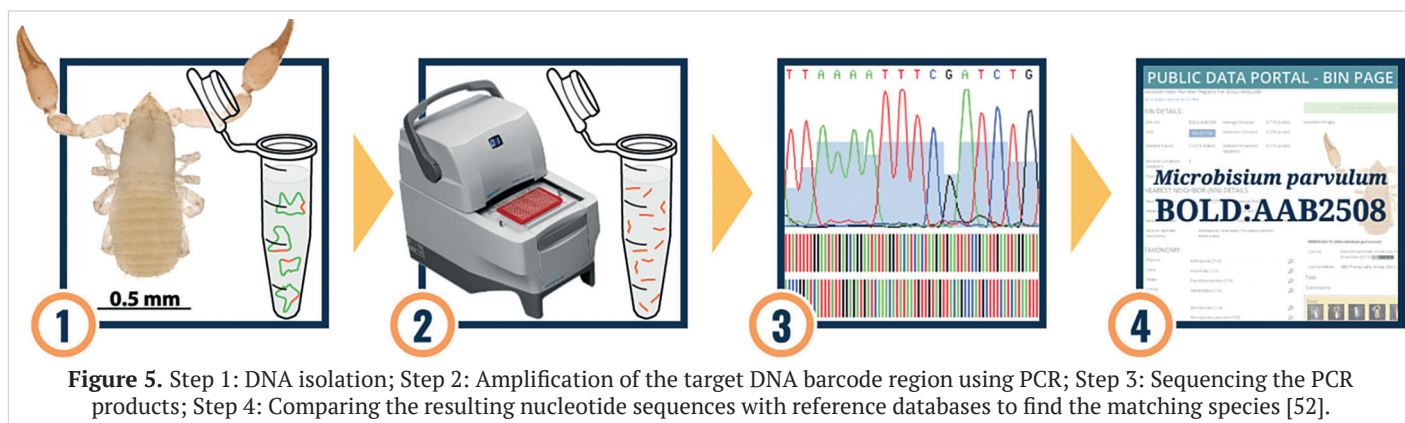
### 2.5. DNA barcoding

The main idea of DNA barcoding resides in the fact that some short DNA region can play a role of a marker that allows definitive identification (or almost definitive, as there is intraspecies variability) of species origin of an organism similarly to the work of a barcode on a label, which is read by a scanner upon payment for goods [51,52].

Figure 5 presents the technology of DNA barcoding. As can be seen from the scheme, the DNA barcoding basis includes two methods of molecular diagnostics: first of all, the polymerase chain reaction, which due to the DNA ability to replicate allows accumulation (amplification) of a chosen fragment of its molecule in quantities suitable for further analysis; secondly, the method for determination of nucleotide sequence in DNA molecules (its sequencing) [51,52].

In 2003, Hebert et al. [42] introduced the term DNA barcoding for the first time in the paper “Biological identifications through DNA barcodes” and proposed to use the 648 bp portion of the mitochondrial cytochrome c oxidase I (COI) gene as a marker DNA sequence to create the global system for animal bio-identification [53,42,54]. Hebert’s study caused mixed reactions varied from enthusiasm, especially in ecologists [55] to criticism, mainly regarding identification of closely related species using single gene [56,57,58,16]. However, the DNA barcoding method is long-established; its advantages and limitations became obvious [53].

Different tasks are accomplished using the DNA barcoding technology, for example, identification of a plant only by its leaves when its flowers and fruits are unavailable; identification of insect larvae, which have less diagnostic traits than adult individuals; determination of an animal diet by the stomach content, saliva or feces [56,60].



**Figure 5.** Step 1: DNA isolation; Step 2: Amplification of the target DNA barcode region using PCR; Step 3: Sequencing the PCR products; Step 4: Comparing the resulting nucleotide sequences with reference databases to find the matching species [52].

DNA barcoding is also used as a molecular tool for determination of food product mislabeling and revelation of species adulteration [61].

The DNA barcoding technology is inextricably linked with Sanger sequencing — a gold standard of DNA typing.

Reading a nucleotide sequence began with the development of the method for sequencing RNA obtained from the DNA template using RNA polymerase. In 1976, therefore, the sequence of the most part of the genome of DNA virus SV40, which length is more than 5000 base pairs, was determined [62,63]. Then, the methods for direct DNA sequencing were developed.

In 1975, F. Sanger and A. Coulson developed the method of direct enzymatic DNA sequencing that is also called plus and minus method. A fragment of single stranded DNA served as a template in the polymerase chain reaction and synthetic complementary sequences or short DNA regions obtained by the action of restriction endonucleases were used as primers [64].

The method consisted of two steps. First, under limiting conditions, the polymerase reaction was performed in the presence of all four dNTP types (one of them was labeled on the alpha position of phosphate), obtaining, in the end, a set of products of incomplete copying of a template fragment. The mixture was then purified from unbound deoxynucleoside triphosphates and divided into eight parts. After that, in the plus system, four reactions were carried out in the presence of each type of nucleotides and in the minus system in the absence of each of them. As a result, in the minus system termination occurred before dNTP of the given type, and in the plus system after it. Eight samples obtained in such a way were separated by electrophoresis, the signal was read off and the sequence of the initial DNA was determined. Using this method, phage  $\phi$ X174 short DNA, which consisted of 5386 bp, was sequenced [64,65].

Genome investigations enable solving many applied and fundamental tasks. Using these methods, new drugs and products have been developed; they also allow penetrating the long human history or understand the cause of mass extinctions of species [66].

Due to these projects, the international genome base NCBI was formed. With its use, it is possible to select nucleotide sequences of interest and develop diagnostic test-systems on their basis. Since then, the PCR method have gradually come into routine laboratory practice and ceased to belong only to the fundamental science.

**2.6. Sanger sequencing**

Sanger sequencing allows reading off sequences of up to 1000 base pairs and is used for small fragments of genome/genes. In particular, it is used for sequencing individual genome regions to analyze mutations and polymorphisms; identify viruses and organisms (bacteria, rickettsia, plants, fungi and animals); validate data obtained on the platforms of next generation sequencing

(NGS); microsatellite analysis; analysis of deletions and insertions (small and long) [66].

Since 2015, Sanger sequencing has also been used for identification of fish species composition. From July 1, 2018, the interstate standard GOST 34106–2017 came into force as a national standard of the Russian Federation. This standard regulates the method for sequencing the fragments of the mitochondrial genome of animals and fish to determine species origin in one-component products. The essence of the method described in GOST consists in determination of the nucleotide sequences of the mitochondrial genome region of different animal and fish species and their comparison with known sequences to identify their species origin. The analysis by this method includes: DNA extraction and purification; PCR with primer pairs flanking the site of the mitochondrial genome in the region of cytB gene; detection of PCR products by the method of electrophoresis in agarose gel to reveal the specific band of amplified DNA and assessment of the concentration of the PCR product; sequencing the PCR product purified from unbound primers and dNTP by the di-deoxynucleotide method with fluorescent dyes; identification of a nucleotide sequence by separating products of the sequencing reaction purified from the excess of dNTP, fluorescently labeled ddNTPs, primer and salts, by the method of capillary electrophoresis and comparison of obtained nucleotide sequence of the genome fragment extracted from the analyzed sample with known sequences from databases for its identification [67].

At present, sequencers are produced also in Russia; the best known model is the genetic analyzer “Nanofor 05”.

The main advantages of Sanger sequencing are high accuracy of reading (many PCR methods are validated by its use) and low primary cost of analysis. However, Sanger sequencing has a serious limitation — it is impossible to carry out species identification in multicomponent products by this method.

Only when using multiplex next generation sequencing (NGS), it is possible to identify complete species composition of multicomponent products in one reaction.

**2.7. Next generation sequencing**

Next generation sequencing began with the principal discovery of the possibility of clonal amplification of fragmented DNA on the solid surface. If in Sanger sequencing, information was read off from one amplified DNA fragment, in next generation sequencing, there is parallel reading from tens of thousands of amplified fragments, where every fragment is cloned and read in its own cell.

Contrary to Sanger sequencing, the NGS methods are used for deep (multiple) read of genetic material which is necessary, for example, for re-sequencing and assembly of novel genomes (*de novo*), transcriptomic and epigenomic investigations. In addition, next generation sequencing (NGS) is much more efficient allowing reading millions and even billions of short fragments.



Such growth in performance led to a possibility to simultaneously determine sequences of tens of genomes (depending on their sizes) in a single run [66].

Next generation sequencing gave an opportunity to assess metagenome of mixed microbial populations and reveal previously unknown and uncultivable microbial forms. This became possible with the development of the algorithms for assembly and analysis of genomes, software and high-performance workstations.

New mathematical and information technologies allow genomics to develop quicker and use more complex algorithms. These algorithms can include simultaneously several applications and programs, and enable working with very large volume of data [66].

When analyzing whole genomes, two approaches are used: analysis by alignment against the reference genome or re-sequencing, or genome assembly from zero or sequencing de novo.

2.8. Basic next generation sequencing methods

The first technology of next generation sequencing is the 454-sequencing technology (high-throughput DNA pyrosequencing or pyrophosphate sequencing) developed by 454 Life Sciences. On the basis of this technology, the methodical recommendations on species identification of fish and fish products based on sequencing of amplified DNA fragments were approved and presented in Russia in 2015 [68]. The amplified DNA fragment with a length of 350–500 base pairs is obtained in the course of two PCR rounds using fusion primers consisted of several parts: the A and B adaptors, four-nucleotide ‘key’ sequence as well as the sequences specific to a target gene. Optionally, the multiplex identifier (MID) sequence is included in the primer sequence. Then, the obtained amplification products are subjected to emulsion PCR (ePCR), during which amplified fragments of the DNA library and special spherical beads are incorporated into water droplets of emulsion, which are amplification microreactors. Clonal parallel amplification in emulsion droplets ensures obtaining millions of copies of each fragment from the DNA library on each individual bead. Using centrifugation, beads are loaded onto a plate containing several hundreds of thousands of microscopic wells. Each bead gets into an individual well of the plate; then, the plate is placed into an instrument along with reagents and sequencing begins. During the work of the instrument, nucleotides are successively flowed through the plate wells. Incorporation of a nucleotide complementary to the template on each individual bead leads to a chemiluminescent signal in a certain well that is recorded by a camera of the instrument.

This technology allows reading simultaneously several genome regions, which makes it possible to identify fish species in the composition of multicomponent products [68].

The obtained nucleotide sequence of the specific DNA fragment is analyzed in the international database NCBI [69].

Figure 6 presents a fragment of a file with the results of analysis by 454 sequencing; the file contains a pool of consensus reads in the FASTA format marked by the identifier clusterid=X. The obtained sequences are copied into the memory buffer and entered into the dialog window “Nucleotide BLAST” of the database NCBI (Figure 7). After that, the analysis program starts and the result window appears showing reference sequences that match those under investigation in the descending order of matching (Figure 8).

2.9. Ion semiconductor sequencing technology

The technology was developed by Ion Torrent Systems, Inc and was launched in 2010 [70]. The principle of the method resides in recording a released hydrogen ion by a ion sensor ISFET.

```

>JCY5JSI01C3T4I length=456 clusterid=1 members=3239
GCCTCTGCCCAAAATGGTACGGCCGCGTATTTTACCGTGCAGGATAGCGCAATCACT
TGCTCTTTAAATGAAGACCTGTATGAAATGGCATCACGAGGGCTTAGCTGTCTCTCTCC
AAGTCAATGAAATTTGATCTGCCCGTGCAGAGCGGACATAAAACACATAAGACGAGAAGA
CCCTATGGAGCTTTAGACACCAGGCAGATCACGTCAAGTAACCTTGAATTAACAAGTAAA
AAGCAGTAGTACCCTTAGCCATATGTCTTTGGTTGGGGCGGACCGGGGGGAAACAAAAG
CCCCATGTGGACTGGGGGCACTGCCCCACACCAAGTACCAACTCTTAAGTACCAGA
ATTTCTGACCAAAAATGATCCGGCATCACGCCGATCAACGGACCGAGTTACCTTAGGGAT
AACCAGCGCAATCCTGAGAGCTTCCAACCACTTCTT
>JCY5JSI01ALM4T length=511 clusterid=2 members=959
AAGAAGTGGTTGGAAGCTCTCACATCAGGCTGCAAGTAGTGTGCAAAATTCGGGTATG
TTATTTATCTTTGTGAAAACGTAGAAAATGATGTTATATATCTGTGTGAAGCATAAAA
TTAATCAATCAATCACTCAATGCAAAAATGCAAAATAGTACATGCAAAAACACAGATATTA
AAAAAATTCRAAAAATATACCTGTAGTAGAGCATGTGGGTA AAAAGTTAATTTGTTA
TAACCTAAAAAAATTTGCATTAATGTGACTCATTAGATTTGAGCCAGTACCACATAGACA
TTTTAAGTAATAATGACTTTTCATTTGACTGAGTTCAACTTACTAGAAGTATTTGAGTTAA
AAATGCCTTTGGTTTACAGTGCAAAAGTGGCCACCAAAATGGTCATTTGTTCTCCCT
CCTTTCTTACCTTGATGAGCAGATTGATGCCCACTTTGAGGTCGCTGAGCTTCTCAGAG
ATCTGGTTGGAGTTTCTGACCATTAGGCTGA
>JCY5JSI01AC9WX length=503 clusterid=3 members=460
GCCTCTGCCCAAAATGGTACGGCTAATGGTACAGAACTCCAACCGATCTCTGAGAAGCT
CAGCGACCTCAAAGTGGGCATCAATCTGCTCATCAAGTAAAGAAAGGGGAGAACAAAT
GACCATTTTGGTGGCCGACTTTGGTGCCTGTAAACCAAGGCAATTTTAACTCAATA
CTCTAGTAAGTTGAACTCAGTCAATGAAAAGTCAATTTACTTAAATGTCTATGTTGGT
ACTGGCTCAATCTAATGAGTCACATTAATGCAATTTTTTAAAGTTATAACRAATTAAC
TTTTACCCAGCATGCTCTACTACAGGTATATTTTGGAAATGTTTTTAAATCTGTGTTT
TTGCATGTAATTTGCAATTTTGCATTTGCAATTTGATGATTTGATTAATTTTATGCTTACA
CAGATATATAACATACATTTTCTACGTTTTTCAAAAATGATAAATAACATACCGGAATTT
GCAAACTACTTGCAGGCTGAT
    
```

Figure 6. A file fragment with the results obtained in the course of sequencing of amplified DNA fragments using the 454-sequencing technology (Figure from the authors’ archive)

It is released when an introduced nucleotide is complementary to the single-stranded fragment of amplified DNA. The reaction and recording of the event occur on a special semiconductor chip with different capacities. In contrary to other systems, this allows scaling performed investigations selecting a chip with a corresponding capacity without buying a more efficient sequencer.

However, compared to the 454 Life Sciences technology (pyrosequencing), the semiconductor technology does not yet allow reading fragments more than 400 bp. Although, in 2018, the company released chips with the possibility to read fragments of up to 600 bp, this immediately doubles the cost of analysis compared to the standard 400 bp reads.

Also, a huge plus of this technology is a possibility of full automation of the sequencing process. A minus, probably, is closeness of this technology compared, for example, with Illumina.

At present, methods based on next generation sequencing have begun to be used in fish species identification as only they are capable of complete revelation of species composition in multicomponent fish products. The only disadvantage of these methods is the price of analysis.

3. Conclusion

Methods based on DNA analysis are often used for fish and fish product species identification. Therefore, they can be used in controlling fish products for correspondence to the species composition indicated on a label.

The most promising are PCR-based methods, which allow identifying meat of different, even closely related fish species. The potential for using the PCR method proceeds from the possibility to identify meat of different, even closely related fish species. This identification can be performed on the biological material obtained from raw tissue, as well as on muscles subjected to different technological treatments. Relative simplicity and rapidness of analyses allow suggesting that these methods consisted in DNA analysis will find wide application in the future for food quality control.

However, despite undoubted advantages of PCR-based methods, we should not overlook their disadvantages.

Generally speaking, SSCP analysis is not recommended to use for the identification of products consisting of several species

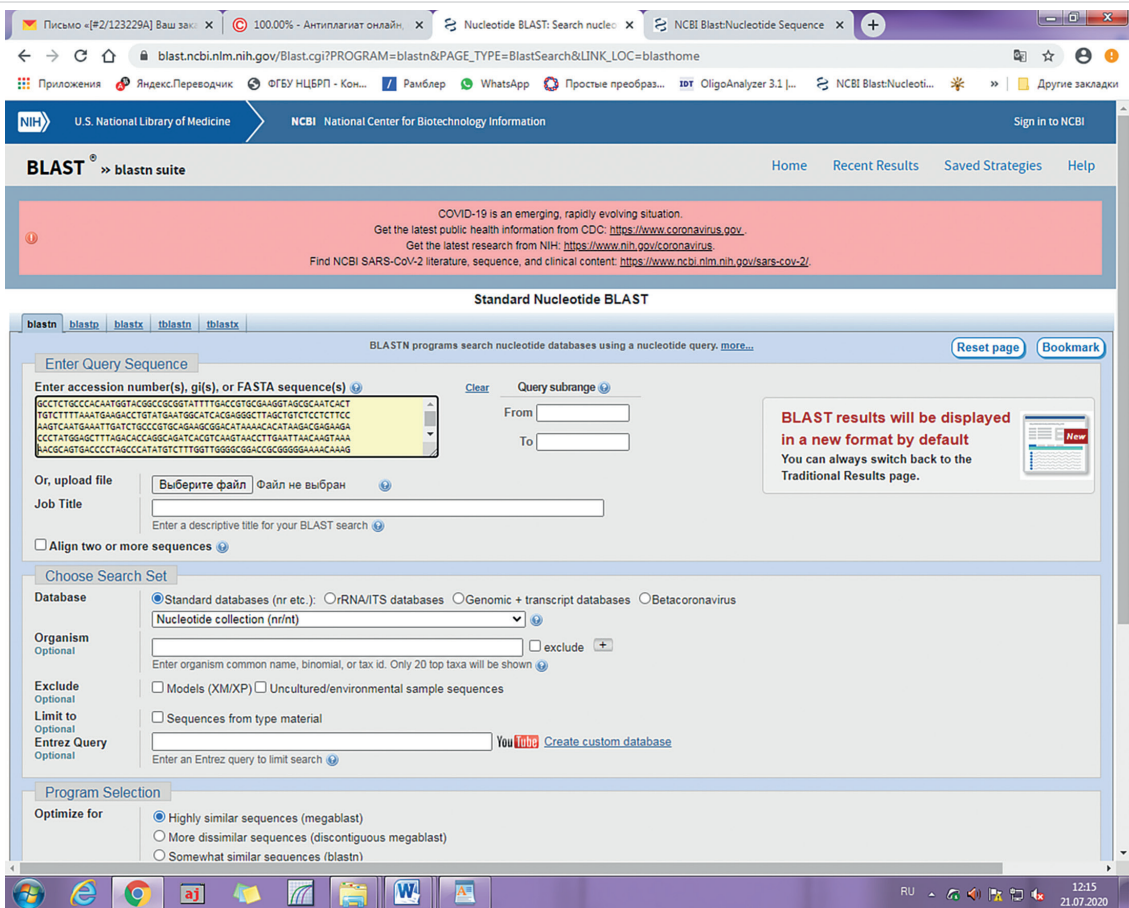


Figure 7. The dialog window “Nucleotide BLAST” of the NCBI database [69]

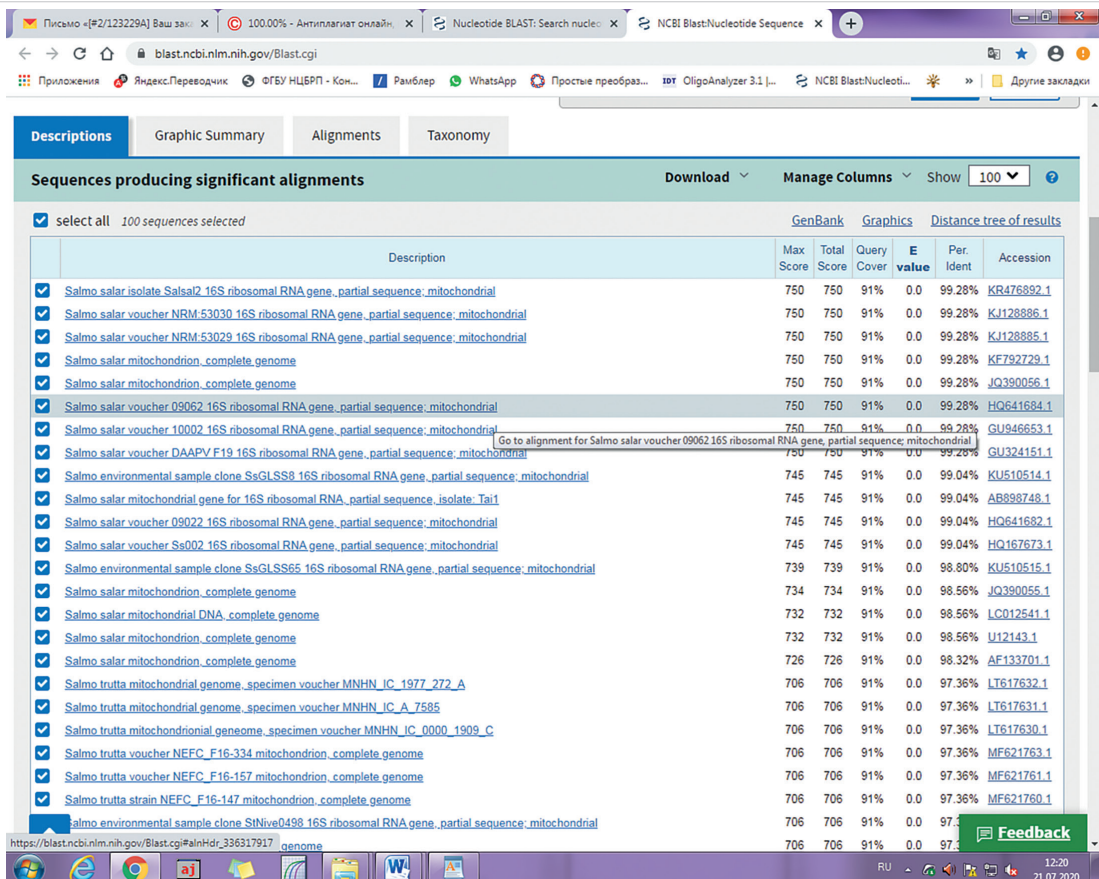


Figure 8. The window with the results of the analysis of the nucleotide sequence obtained in the course of sequencing of amplified DNA fragments using the 454 sequencing technology [69]



of fish and in products subjected to intense heating (sterilization). In addition, a prerequisite for species identification using the SSCP analysis is the presence of reference samples and test samples on the same electrophoretic gel. The repeatability of obtained results in the SSCP analysis is influenced by analysis conditions, that is, a temperature and concentration of used reagents (for example, buffers).

According to several researchers, the PCR-RFLP method is also not recommended for verification of species composition in fish mixtures, as obtained results not always reflect the true composition of a mixture. In case of the PCR-RFLP method, there is a risk of incomplete restriction of a region or intraspecies differences, which can facilitate deletion or extension of restriction sites.

Species-specific PCR in the conventional variant envisages the use of the step of electrophoresis and gel staining to detect the amplification results, which can lead to contamination and, consequently, to unreliable results. In this case, the modification of conventional PCR — real-time PCR — can be considered the most promising method.

Real-time PCR is the most common technology to use for species identification. Continuous measurement of fluorescence allows eliminating the stages (electrophoresis and gel staining)

that usually are necessary to perform after PCR. In addition, in case of real-time PCR, the potential risks of contamination are significantly reduced as samples with the reaction mixture remain to be sealed throughout analysis. At present, real-time PCR is the most common technology to use for species identification in all product types ranging from raw materials to thermally processed products including products with the multicomponent composition from different fish types.

The DNA barcoding method with the use of next generation sequencing is also quite promising. It can detect the presence of all fish species contained in the test sample over one analysis, and the ability to carry out analyses of any fish products makes this method universal. The main disadvantage of the DNA barcoding method using next generation sequencing is its high cost, which is made up from high cost of equipment for analysis, reagents and diagnostic test-systems.

It is necessary to emphasize that PCR methods are only those methods that allow determining product species composition; however, complex quality and safety assessment of fish products also includes analyses by other methods, which simultaneously can confirm or do not confirm correspondence of the producer declaration to what consumer purchased.

## REFERENCES

- Ferrito, V., A. Pappalardo, A.M. (2017). Seafood species identification by DNA barcoding, a molecular tool for food traceability. *Biodiversity Journal*, 8(1), 65–72.
- Carrera, M., Cañas, B., Gallardo, J.M. (2013). Fish Authentication. Chapter in the book *Proteomics in Foods*, 205–222. [https://doi.org/10.1007/978-1-4614-5626-1\\_12](https://doi.org/10.1007/978-1-4614-5626-1_12), ISBN 978-1-4614-5625-4
- Jacquet, J.L., Pauly, D. (2008). Trade Secrets: Renaming and Mislabeling of Seafood. *Marine Policy*, 32(3), 309–318. <https://doi.org/10.1016/j.marpol.2007.06.007>
- von den Heyden, S., Barendse, J., Seebregts, A.J., Mattee, C.A. (2010). Misleading the masses: Detection of Mislabeled and Substituted Frozen Fish Products in South Africa. *ICES Journal of Marine Science*, 67(1), 176–185. <https://doi.org/10.1093/icesjms/fsp222>
- Miller, D.D., Mariani, S. (2010). Smoke, mirrors and mislabeled cod: Poor transparency in the European seafood industry. *Frontiers in Ecology and the Environment*, 8(10), 517–521. <https://doi.org/10.1890/090212>
- Chandrika, M., Maimunah, M., Zainon, M.N., Son, R. (2010). Identification of the species origin of commercially available processed food products by mitochondrial DNA analysis. *International Food Research Journal*, 17(4), 867–876.
- Patel, R.R., Thakkar, N.J., Shah, P.D., Mankodi, P.C. (2018). Species confirmation and evaluation of nutritive values of frozen fish products. *International Journal of Food Science and Nutrition*, 3(2), 58–63.
- Regulation (EU) No 1169/2011 of 25 October 2011 on the provision of food information to consumers, amending Regulations (EC) No. 1924/2006 and (EC) No. 1925/2006 of the European Parliament and of the Council, and repealing Commission Directive 87/250/EEC, Council Directive 90/496/EEC, Commission Directive 1999/10/EC, Directive 2000/13/EC of the European Parliament and of the Council, Commission Directives 2002/67/EC and 2008/5/EC and Commission Regulation (EC) No 608/2004. Official Journal of the European Union, L304, 18–63.
- Technical regulations of the Customs Union TR CU022/2011 “Food products in terms of its labeling” (approved by the decision of the customs Union Commission of December 9, 2011 № 880) (as amended on September 14, 2018). Moscow. — 2011. (in Russian)
- Esakova, N. V. Pampura, A.N., Varlamov, E.E., Okuneva, T.S. (2017). Clinical and immunological features of anaphylaxis to fish in children. *Experimental and clinical gastroenterology journal*, 1(137), 78–82. (In Russian)
- Sharp, M. F., Lopata, A. L. (2014). Fish allergy: In review. *Clinical Reviews in Allergy and Immunology*, 46(3), 258–271. <https://doi.org/10.1007/s12016-013-8363-1>
- Pampura, A.N., Varlamov E. E. (2019). The clinical significance of food animal allergens. *Russian Journal of Allergy*, 16(1–1), 29–35. (In Russian)
- Telmo J. R. Fernandes, Joana Costa, M. Beatriz P. P. Oliveira, Isabel Mafra, Telmo J. R. Fernandes, T.J.R., Costa, J., Oliveira, M.B. P.P., Mafra, I. (2018). Exploiting 16S rRNA gene for the detection and quantification of fish as a potential allergenic food: A comparison of two real-time PCR approaches. *Food Chemistry*, 45, 1034–1041. <https://doi.org/10.1016/j.foodchem.2017.11.068>
- Namazatova-Baranova, L.S. (2011). Allergy in children: from theory to practice. Moscow: The Union of pediatricians of Russia. — 668 p. ISBN 978-5-904753-06-1 (In Russian)
- Kuprina, E.E. (2015). Identification of commercial hydrobionts by ichthyological and instrumental methods. St. Petersburg: ITMO University. — 110 p. (In Russian)
- Teletchea, F. (2009). Molecular identification methods of fish species: re-assessment and possible applications. *Reviews in Fish Biology and Fisheries*, 19(3), 265–295. <https://doi.org/10.1007/s11160-009-9107-4>
- Boidya, P., Haque, W., Rahman, M.M. (2015). Molecular identification and phylogenetic assessment of some marine catfishes of the bay of bengal. *International Journal of Pure and Applied Zoology*, 3(4), 279–286.
- Pozdnyakovskiy, V.M., Ryazanova, O.A., Kalenik, T.K., Datsun, V.M. (2007). Inspection of fish, fish products and non-fish objects of aquatic harvesting. Quality and safety. Novosibirsk: Siberian University Publishing House. — 306 p. ISBN 5-94087-041-4 (In Russian)
- Minchenko, A.G., Dudareva, N.A. (1990). The mitochondrial genome. Novosibirsk: Nauka. — 194 p. ISBN 5-02-029553-1 (In Russian)
- Lindah, T. (1995). Instability and decay of the primary structure of DNA. *Nature*, 362(6422), 709–715. <https://doi.org/10.1038/362709a0>
- Modica-Napolitano, J.S., Singh, K. (2002). Mitochondria as targets for detection and treatment of cancer. *Expert Reviews in Molecular Medicine*, 4(9), 1–19. <https://doi.org/10.1017/s1462399402004453>
- Petros, J.A., Baumann, A.K., Ruiz-Pesini, E., Amin, M.B., Sun, C.Q., Hall, J., Lim, S.-D., Issa, M.M., Flanders, W.D., Hosseini, S.H., Marshall, F.F., Wallace, D.C. (2005). mtDNA mutations increase tumorigenicity in prostate cancer. *Proceedings of the National Academy of Sciences of the United States of America*, 102(3), 719–724. <https://doi.org/10.1073/pnas.0408894102>
- Richter, C. (1987). Biophysical consequence of lipid peroxidation in membranes. *Chemistry and Physics of Lipids*, 44(2–4), 175–189. [https://doi.org/10.1016/0009-3084\(87\)90049-1](https://doi.org/10.1016/0009-3084(87)90049-1)
- Nurgaliev, A. Kh., Karunas, A.S., Khusainova, R.I., Khidiyatova, I.M., Akhmetova, V.L., Valiev, R.R., Nadyrshina, D.D., Mustafin, R.N., Murzabaeva, S. Sh., Khusnutdinovs, E.K. (2013). Molecular-genetic methods for studying human hereditary diseases. Ufa: Bashkir State University. — 102 p. (In Russian)
- Sunnucks, P., Wilson, A.C.C., Beheregaray, L.B., Zenger, French, K.J., Taylor, A.C. (2000). SSCP is not so difficult: the application and utility of single-stranded conformation polymorphism in evolutionary biology and molecular ecology. *Molecular Ecology*, 9(11), 1699–1710. <https://doi.org/10.1046/j.1365-294x.2000.01084.x>
- Konstantinos, K. V., Panagiotis, P., Antonios, V. T., Agelos, P., Argiris, N.V. (2008). PCR-SSCP: A Method for the Molecular Analysis of Genetic Diseases. *Molecular Biotechnology*, 38(2), 155–163. <https://doi.org/10.1007/s12033-007-9006-7>
- Weder, J., Rehbein, H., Kaiszer, K.-P. (2001). On the specificity of tuna-directed primers in PCR-SSCP analysis of fish and meat. *European Food Research and Technology*, 215(2), 139–144. <https://doi.org/10.1007/s002170100339>
- Ludwig, A. Identification of Acipenseriformes species in Trade. [Electronic resource: [http://awsassets.panda.org/downloads/24\\_2008\\_identification\\_acipenseriformes\\_in\\_trade\\_jucn\\_int.pdf](http://awsassets.panda.org/downloads/24_2008_identification_acipenseriformes_in_trade_jucn_int.pdf) Access date 25.05.2020]
- Spychaj, A., Mozdziak, P.E., Pospiech, E. (2009). PCR methods in meat species identification as a tool for the verification of regional and traditional meat products. *Acta Scientiarum Polonorum Technologia Alimentaria*, 8(2), 5–20.



30. Bobo, L.D. (1993). PCR Detection of Chlamydia trachomatis. Diagnostic. Molecular Microbiology. Principles and Applications. Washington: ASM Press. — P. 235–241.
31. Birstein, V.J., Doukakis, P., Sorkin, B., DeSalle, R. (1998). Population aggregation analysis of three caviar-producing species of sturgeons and implications for the species identification of black caviar. *Conservation Biology*, 12(4), 766–775. <https://doi.org/10.1046/j.1523-1739.1998.97081.x>
32. Doukakis, P., Birstein, V. J., Ruban, G. I., DeSalle, R. (1999). Molecular genetic analysis among subspecies of two Eurasian sturgeon species, *Acipenser baerii* and *A. stellatus*. *Molecular Ecology*, 8(s1), S117–S127. <https://doi.org/10.1046/j.1365-294X.1999.00816.x>
33. Coşier, V. (2019). Multiplex PCR assay for detection and identification of animal species in the meat products. *Lucrari Stiintifice. Seria Zootehnie*, 72, 215–220.
34. Michelini, E., Cevenini, L., Mezzanotte, L., Simoni, P., Baraldini, M., De Laude, L., Roda, A. (2007). One-step triplex-polymerase chain reaction assay for the authentication of yellowfin (*Thunnus albacares*), bigeye (*Thunnus obesus*), and skipjack (*Katsuwonus pelamis*) tuna DNA from fresh, frozen, and canned tuna samples. *Journal of Agricultural and Food Chemistry*, 55(19), 7638–7647. <https://doi.org/10.1021/jf070902k>
35. Higuchi, R., Fockler, C., Dollinger, G., Watson, R. (1993). Kinetic PCR Analysis: Real-time monitoring of DNA amplification reaction. *Nature Biotechnology*, 11(9), 1026–1030. <https://doi.org/10.1038/nbt0993-1026>
36. Heid, C.A., Stevens, J., Livak, K.J., Williams, P.M. (1996). Real-time quantitative PCR. *Genome Research*, 6(10), 986–994. <https://doi.org/10.1101/gr.6.10.986>
37. Ekimov, A.N., Shipulin, G.A., Bochkarev, E.G., Rumin, D. V. Real-Time PCR. [Electronic resource: <https://www.interlabservice.ru/catalog/faq/?id=3422> Access date 11.05.2020] (In Russian)
38. Tyagi, S., Kramer, F.R. (1996). Molecular beacons: probes that fluoresce upon hybridization. *Nature Biotechnology*, 14(5), 303–308. <https://doi.org/10.1038/nbt0396-303>
39. Hird, H. J., Hold, G. L., Chisholm, J., Reece, P., Russell, V.J., Brown, J., Goodier, R., MacArthur, R. (2005). Development of a method for the quantification of haddock (*Melanogrammus aeglefinus*) in commercial products using real-time PCR. *European Food Research and Technology*, 220(5–6), 663–637. <https://doi.org/10.1007/s00217-004-1050-y>
40. Sánchez, A., Quinteiro, J., Vázquez, J.A., Perez-Martín, R.I., Sotelo, C.G. A. (2019). Comparison of real-time PCR methods for quantification of European hake (*Merluccius merluccius*) in processed food samples. *Food Chemistry*, 272, 279–285. <https://doi.org/10.1016/j.foodchem.2018.08.031>
41. Tyulkin, S.V., Vafin, R.R., Muratova, A.V., Khatypov, I.I., Zagidullin, L.R., Rachkova, E.N., Akhmetov, T.M., Ravilov, R.K. (2015). Development of a method for PCR-RFLP on the example of *dgat1* gene in cattle. *Fundamental research*, 2–17, 3773–3775. (in Russian)
42. Hebert, P.D.N., Cywinska, A., Ball, S.L., deWaard, J.R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1512), 313–321. <https://doi.org/10.1098/rspb.2002.2218>
43. Gil, L.A. (2007). PCR – based methods for fish and fishery products authentication. *Trends in Food Science and Technology*, 18(11), 558–566. <https://doi.org/10.1016/j.tifs.2007.04.016>
44. Hold, G.L., Russell, V.J., Pryde, S.E., Rehbein, H., Quinteiro, J., Rey-Mendez, M., Sotelo, C.G., Pérez-Martín, R.I., Santos, A.T., Rosa, C. (2001). Validation of a PCR-RFLP based method for the identification of salmon species in food products. *European Food Research and Technology*, 212(3), 385–389. <https://doi.org/10.1007/s002170000237>
45. Russell, V.J., Hold, G.L., Pryde, S.E., Rehbein, H., Quinterio, J., Rey-Mendez, M., Sotelo, C.G., Pérez-Martín, R.I., Santos, A., Rosa, C. (2000). Use of restriction fragment length polymorphism (RFLP) to distinguish between salmon species. *Journal of Agricultural and Food Chemistry*, 48(6), 2184–2188. <https://doi.org/10.1021/jf991213e>
46. Lin, W.-F., Hwang, D.-F. (2007). Application of PCR-RFLP analysis on species identification of canned tuna. *Food Control*, 18(9), 1050–1057. <https://doi.org/10.1016/j.foodcont.2006.07.001>
47. Wu, Ya-J., Hsieh, C.-H., Chen, H.-M., Hwang, D.-F. (2008). Identification of six common species of processed filefish using cytochrome b gene sequence and PCR-RFLP analysis. *The raffles bulletin of zoology*, 19, 151–158.
48. Mueller, S., Ravi, H., Novoradovskaya, N., Kincaid, R., Chee, Y.-L. (2011). Enhanced fish species identification by PCR-RFLP using the 2100 Bio-analyzer system. *International Food Research Journal*, 18(3), 1209–1215.
49. Dooley, J. J., Sage, H. D., Brown, H. M., Garrett, S. D. (2004). Improved fish species identification by use of lab-on-a-chip technology. *Food Control*, 16(7), 601–607. <https://doi.org/10.1016/j.foodcont.2004.06.022>
50. Formosa, R., Ravi, H., Happe, S., Huffman, D., Novoradovskaya, N., Kincaid, R., Garrett, S. (2010). DNA-based Fish Species Identification Protocol. *Journal of Visualized Experiments*, 38, e1871. <https://doi.org/10.3791/1871>
51. Zakharov, I.A., Shaikovich, E.V., Ivshin, N.V. (2007). DNA-barcoding in entomology. *Priroda*, 9(1105), 3–9. (In Russian)
52. What is DNA Barcoding? [Electronic resource: [www.ibol.org](http://www.ibol.org). Access date 12.06.2020]
53. Shekhovtsov, S.V., Shekhovtsova, I.N., Peltek S. E. (2019). DNA barcoding: methods and approaches. *Uspehi sovremennoj biologii*, 139(3), 211–220. <https://doi.org/10.1134/S0042132419030074> (In Russian)
54. Ward, R.D., Hanner, R., Hebert, P.D.N. (2009). The Campaign to DNA Barcode All Fishes, FISH-BOL. *Journal of Fish Biology*, 74(2), 329–356. <https://doi.org/10.1111/j.1095-8649.2008.02080.x>
55. Janzen, D.H. (2004) Now is the time. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359(1444), 731–732. <https://doi.org/10.1098/rstb.2003.1444>
56. Lipscomb, D., Platnick, N., Wheeler, Q. (2003). The intellectual content of taxonomy: a comment on DNA taxonomy. *Trends in Ecology and Evolution*, 18(2), 65–66. [https://doi.org/10.1016/s0169-5347\(02\)00060-5](https://doi.org/10.1016/s0169-5347(02)00060-5)
57. Soininen E. M., Valentini A., Coissac E., Miquel C., Gielly L., Brochmann C., Brysting A. K., Sønstebø J. H., Ims R. A., Yoccoz N. G., Taberlet P. (2009). Analysing diet of small herbivores: the efficiency of DNA barcoding coupled with high-throughput pyrosequencing for deciphering the composition of complex plant mixtures. *Frontiers in zoology*, 6(1), 16. <https://doi.org/10.1186/1742-9994-6-16>
58. Mallet, J., Willmott, K (2003). Taxonomy: renaissance or Tower of Babel? *Trends in Ecology and Evolution*, 18, 57–59. [https://doi.org/10.1016/s0169-5347\(02\)00061-7](https://doi.org/10.1016/s0169-5347(02)00061-7)
59. Moritz, C., Cicero, C. (2004). DNA barcoding: promise and pitfalls. *PLoS Biology*, 2(10), e354. <https://doi.org/10.1371/journal.pbio.0020354>
60. Kress, W. J., Wurdack, K. J., Zimmer, E. A., Weigt, L. A., Janzen, D. H. (2005). Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences of the United States of America*, 102(23), 8369–8374. <https://doi.org/10.1073/pnas.0503123102>
61. Barcaccia, G., Lucchin, M., Cassandro, M. (2015). DNA Barcoding as a Molecular Tool to Track Down Mislabeling and Food Piracy. *Diversity*. 8(4), 2. <https://doi.org/10.3390/d8010002>.
62. Fiers, W., Contreras, R., Haegeman, G., Rogiers, R., Van de Voorde, A., Heuverswyn, H. V., Herreweghe, J.V., Volckaert, G., Ysebaert, M. (1978). Complete nucleotide sequence of SV40 DNA. *Nature*, 273(5658), 113–120. <https://doi.org/10.1038/273113a0>
63. Reddy, V.B., Thimmappaya, B., Dhar, R., Subramania, K.N., Zain, B.S., Pan, J., Ghosh, P.K., Celma, M.L., Weissman, S.M. (1978). The genome of simian virus 40. *Science*, 200(4341), 494–502. <https://doi.org/10.1126/science.205947>
64. Sanger, F., Coulson, A.R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology*, 94(3), 444–448. [https://doi.org/10.1016/0022-2836\(75\)90213-2](https://doi.org/10.1016/0022-2836(75)90213-2)
65. Budilov, A. Methods for deciphering nucleotide sequences of DNA fragments. [Electronic resource: [http://molbiol.ru/protocol/13\\_03.html](http://molbiol.ru/protocol/13_03.html) Access date 20.05.2020] (In Russian)
66. Nedoluzhko, A. (2017). Methods in pictures: sequencing of nucleic acids. [Electronic resource: <https://www.researchgate.net/publication/319058696> Access date 01.06.2020] (in Russian)
67. GOST 34106–2017. “Food and raw materials. Sequencing of the mitochondrial genomes of animals and fish for species identification in single component products”. Moscow: Standartinform. 2017. —21 p. (In Russian)
68. MR № 4.0002–15 “Fish and fish products. Methods for species identification based on sequencing of amplified DNA fragments”. Moscow: VNIIMP. —20 p. (In Russian)
69. Standard Nucleotide BLAST [Electronic resource: [http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE\\_TYPE=BlastSearch&LINK\\_LOC=blasthome](http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome) Access date 10.06.2020]
70. Zubov, V. Next Generation Sequencing Technologies (Draft Roadmap). [Electronic resource: <http://bioinformatics.ru/Misc/genseq-roadmap.html> Access date 01.06.2020] (In Russian)

## AUTHOR INFORMATION

**Tatyana A. Fomina** — candidate of technical sciences, senior research scientist, Laboratory of Molecular Biology and Bioinformatics, V. M. Gorbatov Federal Research Center for Food Systems of RAS, 109316, Moscow, Talalikhina str., 26.

Head of the Department of molecular diagnostic research at the testing reference laboratory, National Centre for Safety of aquatic fisheries products and aquaculture. 129626, Moscow, Grafskiy per., 14-1. Tel.: +7-903-732-74-19. E-mail: fomina1032@yandex.ru

\*corresponding author

ORCID: <https://orcid.org/0000-0001-8848-0632>

**Vladimir Yu. Kornienko** — candidate of biological sciences, senior research scientist, Laboratory of Molecular Biology and Bioinformatics, V. M. Gorbatov Federal Research Center for Food Systems of RAS, 109316, Moscow, Talalikhina str., 26. Tel.: +7-495-676-60-11. E-mail: uniprAIM@yandex.ru

<https://orcid.org/0000-0001-9068-9814>

**Mikhail Yu. Minaev** — candidate of technical sciences, head of Laboratory of Molecular Biology and Bioinformatics, V. M. Gorbatov Federal Research Center for Food Systems of Russian Academy of Sciences. 109316, Moscow, Talalikhina str., 26. Tel.: +7-495-676-60-11. E-mail: m.minaev@fncps.ru

<https://orcid.org/0000-0002-0038-9744>

All authors bear responsibility for the work and presented data.

All authors made an equal contribution to the work.

The authors were equally involved in writing the manuscript and bear the equal responsibility for plagiarism.

The authors declare no conflict of interest.

Received 28.06.2020 Accepted in revised 15.09.2020 Accepted for publication 25.09.2020