ISSN No:-2456-2165

Review of Next Generation Sequencing for Meat and Dairy Products Authentication

N G Kala: UG in Department of biotechnology, RV College of Engineering, Autonomous Institute Affiliated to Visvesvaraya Technological University, Belagavi, Bengaluru, Karnataka, India.

Dr. Lingayya Hiremath : Assistant professor ,Department of biotechnology, RV College of Engineering, Autonomous Institute Affiliated to Visvesvaraya Technological University, Belagavi, Bengaluru, Karnataka, India.

Abstract:- The precise authentication of food products is an essential step in an informed consumer choice. Mainly when it comes to meat and dairy products, special consideration to be given for correct labeling as a result of several reasons such as health concerns, religious beliefs, consumer safety and to promote fair trade. There are many methods which are being followed from ages for species identification and most of them are based on either protein or DNA analysis. Next generation sequencing [NGS], a DNA analysis method has been recognized as one of the most reliable technique for species identification. Due to the untargeted nature of NGS, millions of sequences can be produced from individual sequences using different templates at a time which results in detection of thousands of species in each sample. There is no prerequisite knowledge of which target to look for as universal primers are used instead of taxa-specific primers. Being an emerging technology, new NGS platforms are being developed to revolutionize food authenticity testing. This study focuses on why NGS is recognized as the most powerful technology among other DNA based methods for food analysis and provides an overview of the current range of NGS platforms used for meat and dairy product authentication and the potential future development of this technology.

Keywords:- *Authenticity, Next Generation Sequencing, Ion Torrent, Metabarcoding, Bioinformatics, Pyrosequencing.*

I. INTRODUCTION

Consuming genuine meat and dairy products is the root for getting proper nutritional and health benefits for most of the people [1]. FDA refers food fraud as “Economically Motivated Adulteration (EMA),” and it has been recognized as a growing problem on a global scale with wide-ranging health, social, economic and environmental impacts [2]. Nowadays consumers are becoming more aware and informed of ongoing food adulteration problems as media is increasingly accentuating the influences of food fraud. Meat and dairy products, often prone to adulteration practice for economic gain [3]. Revealing contaminated meat and dairy products is an important practice for many reasons, mostly related to health anxieties. Those who hold and act based on religious beliefs consume only certain animal species and therefore they pay special consideration to accurate labeling [4]. As we all are well aware that dairy products are among the most vital sources of fat and proteins for humans and also significant agriculture commodity in terms of world economic value. If there is a substitution or admixture of milk or meat products with different species than those labeled may result in adverse reactions towards allergens of some species [5-7].

Thus, identification of animal species in meat and dairy products is very important in ensuring food safety, quality and consumer protection [8]. Different methods have been employed for identification of meat species, most of them fall under protein and DNA analysis. Lipid and proteins are used for identification of tissues [9]. When we compare DNA and protein analysis for species identification, DNA method is often preferred than proteins. The main explanations offered as the validation are the thermal stability of DNA because of the double helix structure of DNA and the point that DNA subsists in the cells of almost all species [10-12]. In case of most laboratories who are dealing with authenticity issues, they opt for Realtime PCR methods as a first approach to target most common species. The technique involves amplification of primers of particular DNA regions for defined species to be identified. Fluorescence signal is generated if the targeted gene exists in the food product. However, the limitation of Real-time PCR’s fluorescent signal arises if the food product contains multiple species/ingredients. In such cases detection of species cross reactivity may result in false results. From this we can confer that a targeted tactic is not apposite for authenticity concerns as it will only convey a result for the targeted species. Consequently, if the food product contains any additional species in addition to those targeted no information is available by PCR analysis. In such cases the use of untargeted approach comes into play [13, 14].

Such limitations can be overcome using more recent technologies that present themselves as high throughput and low cost approaches in comparison to other methods. One such technology is Next Generation Sequencing [NGS] from which millions of DNA molecules can be sequenced in parallel [15]. This paper provides an overview of how and why NGS can be used for meat and dairy product authentication and which are the platforms available.

ISSN No:-2456-2165

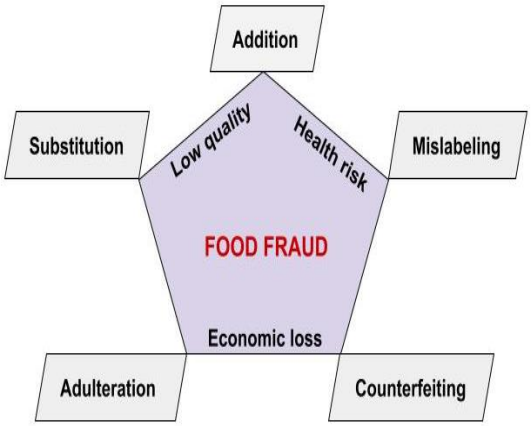

Figure 1 Cause and consequences of food fraud

II. NGS TECHNOLOGY

As previously mentioned, the most common method (Molecular) for verification of species is Real-time PCR and it is limited by the number of targets that can be identified and differentiated at the same time and also requires previous knowledge of the species to be identified. Incorporating Next Generation sequencing into the food sector reforms the food products authenticity testing as it is recognized as the most reliable method for detecting and differentiating species. NGS is an untargeted approach that provides precise identification of thousands of species in each sample using DNA sequencing. There is no need of prerequisite knowledge of which species to search for. So now if any sample is analyzed the purpose of the study is "Which species are present in the sample?" instead of "Are species X, Y or Z present in the sample?" [16].

Next generation sequencing is one complicated method which involves several steps for identifying species in food product. There are many different NGS platforms, each follow their own way of sequencing but the outline of the overall process follow a similar flow. PCR amplicons obtained are sequenced using the NGS approach to get thousands of DNA sequences for every different species that are included in a very sample. The amplicons used are very short (average 100 bp) to maximize success, even in highly processed food products, like food product and animal feed (170ºC, 3 atm). The DNA sequences obtained are compared with internal and/or external DNA databases and species identified. As NGS is an untargeted approach the ultimate result's a listing of all the species that are included in a very sample. As an example, PCR analysis of a meat based product (hamburger, lasagne, pizza) will tell you whether pork is present – yes, or no. With only one test, the DNA Analyzer supported NGS will tell you all the meat species that are present. Adding an NGS analysis of plant DNA, the meat based-product may also be screened for adulteration with material. At < 0.1%, the Limit of Detection (LOD) for NGS is analogous or perhaps less than that of real-time PCR. Additionally, for DNA quantification, NGS also offers advantages over the real-time PCR approach. This is often because universal primers are employed in NGS for the various species during a nutrient, rather than the taxa-specific primers used for PCR which will cause bias during the amplification step. Although NGS isn't yet used routinely for species DNA quantification, this method also will introduce value in terms of its reliability [17].

Next-generation sequencing (NGS) is altering the food industry. There are many misconceptions about NGS technology, as it is a relatively new technology. Best example is that, NGS is often mistaken for whole-genome sequencing (WGS), but in actuality WGS is only one of the applications of NGS. Besides authenticity testing, NGS is also used by food manufacturers and service labs to do much more like pathogen testing, GMO verification, microbiome studies, and persistent analysis.

Next generation sequencing (NGS) tools have reformed the way during which DNA could also be analyzed. These technologies have augmented of several orders of magnitude the sequencing throughput. Their benefit within the species identification rely on the likelihood to combine during one step the generation of species-specific information from the produced short sequence reads with limitless multiplexing potential. Also the possible relative quantification of the detected species by counting reads matching the identical target specific sequence. The large potential and flexibility of NGS is obtained through bioinformatics analysis of the produced sequence data [18, 19]. Among the commercially available bench top NGS platforms which can be used for species identification, it's possible to mention Illumina, Roche 454, and Ion Torrent technologies [18, 20, 21]. Ion Torrent technology relies on the detection of small pH modifications that occur during the sequencing phase which are captured by semiconductor chips accommodating variant reaction micro-wells that produce sequence reads (22). Some reports have tested the Ion Torrent Personal Genome Machine (PGM) sequencing platform for species identification in food products [23-25].

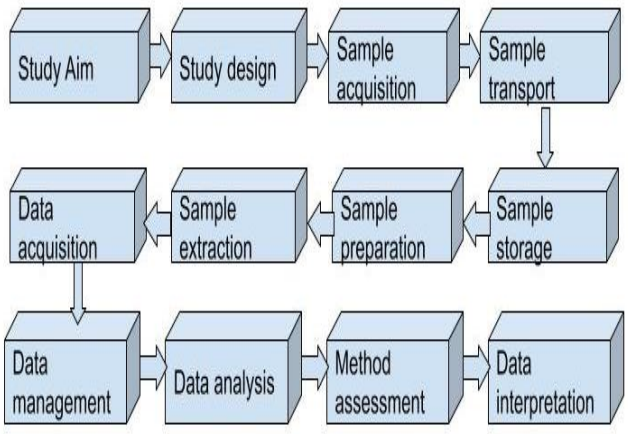

Figure 2: Overview of workflow of non-targeted approach

Illumina has been the most extensively used NGS platform when compared to others for short read sequencing. It employs sequencing by synthesis method, in which the four nucleotides fluoresce at definite and different wavelengths when incorporated by a DNA polymerase. The color signals produced will be sensed by on board cameras and inferred as the DNA sequence of the molecules [26]. One of the earliest

NGS technologies was Roche 454, works on pyrosequencing method. The most recent development of NGS is nanopore sequencing which is also referred as third generation sequencing / fourth generation sequencing. Irrespective of what it is called, nanopore sequencing symbolizes a wholly different method of sequencing. This sequencing technology is quite rapid, handling 450 bases per second and resulting in faster turnaround times and lower costs [27].

III. NGS BASED APPROACHES OR PLATFORMS

Based on variety of different chemistries, many NGS platforms are now widely available for high throughput DNA sequencing, and more sequencing technologies are being developed with bright future [28]. Illumina develops the most widely used NGS technologies which use reversible fluorescent dideoxy terminators for sequencing amplified DNA clusters. Depending on the need of users Illumina offer an array of alternatives which produce a range of different quantities and lengths of DNA reads. For example Illumina iSeq 100, which can develop a maximum of 1.2 billion bases of sequence per run whereas HiSeq X Ten is a set of ten instruments, each producing up to 1.8 trillion bases. The latest releases include NovaSeq 5000 and 6000 which allow the user to regulate output to match their expected requirements. Ion Torrent technology is an alternative platforms that produce short reads. There are currently three Ion Torrent devices, the Proton, and the Gene Studio S5, the Personal Genome Machine [29,30].

A few NGS approaches have been lately applied for species identification, useful for the authentication of food products, validating the powerful potential of these DNA sequencing methods. A research team worked on ion torrent next generation semiconductor based sequencing technology for species identification in complex meat derived products. They used targeted mitochondrial DNA (mtDNA) tactic (intra-species variability) based on dissimilar fragments amplified with three universal primer pairs. Total of 1,363,351 filtered reads were generated by sequencing six libraries. The bioinformatics pipeline relied on the results of the aln algorithm which is known for its performance for species identification discerning closely related sequences. The BLASTN was used for analyses against the GenBank vertebrate mtDNA sequence database. In order to add another computational method and to validate the aln software result they compared the outcome of this algorithm with outputs of BLASTN. The study identified both anticipated and unanticipated meat species in the food products. Accurate meat species ingredient data was obtained successfully for all products. Bubalus bubalis DNA was identified in pork and pork was identified in the kebab [31, 32].

Similar study was conducted on dairy products, in which Ion torrent was used for species identification based on three mtDNA regions obtained from DNA mixtures obtained from dairy products. Dairy products can be used as simple case studies for the application of innovative NGS approaches as they are generally not very intricate in terms of possible species of origin of the milk; Ion Torrent PGM sequencing was executed starting from PCR products. To test the method sequencing reads were derived from 3 libraries that were obtained from pooled amplicons / artificial DNA pools. They tested for two algorithms aln and mem that have unlike protocols for mapping reads and for dairy species aln performed better than mem. The study was done on one goat cheese, one buffalo cheese, one artisanal cheese and two mixed goat and cow milk samples. The sequencing results generated from these five libraries were able to identify all estimated species along with undeclared species. They identified different species mitotypes in the derived products and the presence of human DNA. This study validated the practicality of Next generation sequencing for species verification in dairy yields and its potential application for authentication [31, 33].

DNA barcoding is considered as a gold standard system for species identification which uses short DNA section from specific gene(s). In Recent years, with advance of NGS and also availability of accessible databases such as Genbank, BOLD (Barcode of Life Data Systems) DNA metabarcoding is emerging [34]. A study was done to examine the NGS based DNA metabarcoding to identify avian and mammalian species in mixed products. Sequencing of short segment 16S ribosomal RNA mitochondrial gene was implemented. For most of the animal species, the standard region is the fragment of the mitochondrial gene coding for cytochrome c oxidase subunit I (COI) (approximately 650bp fragment). A high-throughput and ultra-deep sequencing method Illumina HiSeq sequencing was performed. The results showed the existence of mislabeling of meat products [35]. Similarly when mini-barcode system combined with NGS, species identification in highly processed food possibility is increased. Another study demonstrated that 16S rRNA mini-barcoding system provides a new slant in using DNA meta-barcoding for authentication of mixed species-derived food products [41].

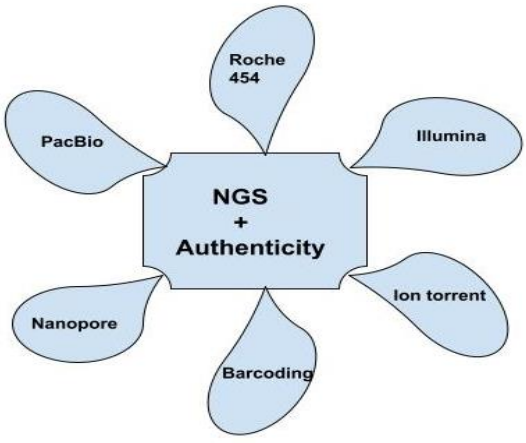

Figure 3: NGS based approaches for food authenticity

When we compare NGS against current ELISA or real-time PCR approaches for species identification, it has 2 momentous advantages over them. Firstly, it is a non-targeted approach that allows recognition of any species that might be present. Furthermore, this methodology is self-validating and data's generated can be easily related with database

sequences [36]. One more NGS platform Pyro-sequencing was first studied by Ortola Vidal and colleagues for food authentication in plant specification [37]. It is one of the rapid NGS technologies that results in quantitative sequence data in the form of a sequence profile [38, 39]. Leatherhead Food Research Company has advanced ASPECT (Adulteration and SPECiation Test). It links highly proficient DNA extraction with non-targeted pyro-sequencing. The method has been used for different assays like meat, fish, dairy, herbs, species and poultry etc. Moreover it is also used for highly processed food products with detection limit of 1% w/w adulterant DNA from the authentic species [40].

IV. CONCLUSIONS

Although there are many techniques for identification of species in meat and dairy products, not all of them are easy, fast, accurate, reasonable, and applicable routine methods in factories, laboratories, control centres. On the other hand the rapidly shifting and growing of Next generation sequencing is changing the way food authenticity is conducted. Moreover, database resources of genetic information are growing faster each year databases. The bioinformatics mining of the reads generated set up a key part of this NGS approach, to be specific when sequences are very identical between species. In the upcoming days, as NGS becomes a well-established method, both methodologies and analysis pipelines should be harmonized among testing laboratories. Technical progresses and augmented race will continue to drive the field towards higher throughput, lower costs, and more user friendly options for analysis. It will also be necessary that the competent testing laboratories become appropriately attributed.

REFERENCES

[1]. Lo YT, Shaw PC. DNA-based techniques for authentication of processed food and food supplements. Food Chem. 2018 Feb 1; 240:767-774. doi: 10.1016/j.foodchem.2017.08.022. Epub 2017 Aug 4. PMID: 28946341.

[2]. David I Ellis, Howbeer Muhamadali, David P Allen, Christopher T Elliott, Royston Goodacre, A flavour of omics approaches for the detection of food fraud, Current Opinion in Food Science, Volume 10, 2016, Pages 7-15, ISSN 2214-7993, https://doi.org/10.1016/j.cofs.2016.07.002.

[3]. Miguel Ángel Sentandreu, Enrique Sentandreu, Authenticity of meat products: Tools against fraud, Food Research International, Volume 60, 2014, Pages 19-29, ISSN 0963-9969, https://doi.org/10.1016/j.foodres.2014.03.030.

[4]. N.Z. Ballin, Authentication of meat and meat products, Meat Science, Volume 86, Issue 3, 2010, Pages 577-587, ISSN 0309-1740, https://doi.org/10.1016/j.meatsci.2010.06.001.

[5]. Høst, A., Jacobsen, H. P., Halken, S., & Holmenlund, D. (1995). The natural history of cow's milk protein allergy/intolerance. European Journal of Clinical Nutrition, 49, Suppl. 1, S13-S18.

[6]. Restani, P., Ballabio, C., Di Lorenzo, C., Tripodi, S., & Fiocchi, A. (2009). Molecular aspects of milk allergens and their role in clinical events. Analytical and Bioanalytical Chemistry, 395, 47-56.

[7]. Umpiérrez, A., Quirce, S., Marañón, F., Cuesta, J., García-Villamuza, Y., Lahoz, C., & Sastre, J. (1999). Allergy to goat and sheep cheese with good tolerance to cow cheese. Clinical and Experimental Allergy, 29, 1064-1068.

[8]. A.H.T. Teh, G.A. Dykes, Encyclopedia of Meat Sciences, 2nd edn. (2014), pp. 265–269

[9]. A.A. Aida, Y.B. Che Man, CMVL Wong, A.R. Raha, R. Son, Analysis of raw meats and fats of pigs using polymerase chain reaction for Halal authentication. Meat Sci. 69, 47–52 (2005)

[10]. Blake, R. D., & Delcourt, S. G. (1998). Thermal stability of DNA. *Nucleic acids research*, *26*(14), 3323–3332. https://doi.org/10.1093/nar/26.14.3323

[11]. V.B. Vallejo-Cordoba, A. González-Córdova, M. MazorraManzano, R. Rodríguez-Ramírez, Capillary electrophoresis for the analysis of meat authenticity. J. Sep. Sci. 28(9–10), 826–836 (2005)

[12]. M. Montowska, E. Pospiech, Authenticity determination of meat and meat products on the protein and DNA basis. Food Rev. Int. 27(1), 84–100 (2010)

[13]. Salihah, N. T., Hossain, M. M., Lubis, H., & Ahmed, M. U. (2016). Trends and advances in food analysis by real-time polymerase chain reaction. *Journal of food science and technology*, *53*(5), 2196–2209. https://doi.org/10.1007/s13197-016-2205-0

[14]. Kralik, P., & Ricchi, M. (2017). A Basic Guide to Real Time PCR in Microbial Diagnostics: Definitions, Parameters, and Everything. *Frontiers in microbiology*, *8*, 108. https://doi.org/10.3389/fmicb.2017.00108

[15]. Coutinho Moraes, D. F., Still, D. W., Lum, M. R., & Hirsch, A. M. (2015). DNA-Based Authentication of Botanicals and Plant-Derived Dietary Supplements: Where Have We Been and Where Are We Going?. Planta medica, 81(9), 687–695. https://doi.org/10.1055/s-0035-1545843

[16]. Slatko, B. E., Gardner, A. F., & Ausubel, F. M. (2018). Overview of Next-Generation Sequencing Technologies. *Current protocols in molecular biology*, *122*(1), e59. https://doi.org/10.1002/cpmb.59

[17]. FOOD AUTHENTICITY TRACEABILITY & SAFETY, UNDERSTANDING DNA-BASED TEST METHODS. NOVEMBER 2018. The Consumer Goods Forum and SGS

[18]. Bertolini, F., Ghionda, M. C., D'Alessandro, E., Geraci, C., Chiofalo, V., & Fontanesi, L. (2015). A next generation semiconductor based sequencing approach for the identification of meat species in DNA mixtures. PLoS One, 10, e0121701.

[19]. Pabinger, S., Dander, A., Fischer, M., Snajder, R., Sperk, M., Efremova, M., et al. (2014). A survey of tools for variant analysis of next-generation genome sequencing data. Briefings in Bioinformatics, 15, 256e278.

[20]. Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: Ten years of next-generation sequencing technologies. Nature Reviews Genetics, 17, 333e351.
[21]. Tillmar, A. O., Dell'Amico, B., Welander, J., & Holmlund, G. (2013). A universal method for species identification of mammals utilizing next generation sequencing for the analysis of DNA mixtures.
[22]. Rothberg, J. M., Hinz, W., Rearick, T. M., Schultz, J., Mileski, W., Davey, M., et al. (2011). An integrated semiconductor device enabling non-optical genome sequencing. Nature, 475, 348e352
[23]. Carvalho, D. C., Palhares, R. M., Drummond, M. G., & Gadanho, M. (2017). Food metagenomics: Next generation sequencing identifies species mixtures and mislabeling within highly processed cod products. Food Control, 80, 183e186
[24]. Giusti, A., Armani, A., & Sotelo, C. G. (2017). Advances in the analysis of complex food matrices: Species identification in surimi-based products using Next Generation Sequencing technologies. PLoS One, 12, e0185586.
[25]. Munoz-Colmenero, M., Martínez, J. L., Roca, A., & Garcia-Vazquez, E. (2017). NGS tools for traceability in candies as high processed food products: Ion Torrent PGM versus conventional PCR-cloning. Food Chemistry, 214, 631e636.
[26]. Drmanac, S., Callow, M., Chen, L., et al., 2020. CoolMPS: advanced massively parallel sequencing using antibodies specific to each natural nucleobase. bioRxiv 953307
[27]. Xiao, T., & Zhou, W. (2020). The third generation sequencing: the advanced approach to genetic diseases. *Translational pediatrics*, *9*(2), 163–173. https://doi.org/10.21037/tp.2020.03.06
[28]. Ansorge, W. J. (2016). Next generation DNA sequencing (II): Techniques, applications. Journal of Next Generation Sequencing & Applications, 01(S1), https://doi.org/10. 4172/2469-9853.s1-005.
[29]. Hodkinson, B. P., & Grice, E. A. (2015). Next-generation sequencing: A review of technologies and tools for wound microbiome research. Advances in Wound Care, 4(1), 50–58. https://doi.org/10.1089/wound.2014.0542
[30]. Song, L., Huang, W., Kang, J., Ren, H., & Ding, K. (2017). Comparison of error correction algorithms for ion torrent PGM data: Application to hepatitis B virus. Scientific Reports, 7, 8106. https://doi.org/10.1038/s41598-017-08139-y
[31]. Anisa Ribani, Giuseppina Schiavo, Valerio Joe Utzeri, Francesca Bertolini, Claudia Geraci, Samuele Bovo, Luca Fontanesi, Application of next generation semiconductor based sequencing for species identification in dairy products, Food Chemistry, Volume 246, 2018, Pages 90-98, ISSN 0308-8146, https://doi.org/10.1016/j.foodchem.2017.11.006.
[32]. Anisa Ribani, Giuseppina Schiavo, Valerio Joe Utzeri, Francesca Bertolini, Claudia Geraci, Samuele Bovo, Luca Fontanesi, Application of next generation semiconductor based sequencing for species identification and analysis of within-species mitotypes useful for authentication of meat derived products, Food Control, Volume 91, 2018, Pages 58-67, ISSN 0956-7135, https://doi.org/10.1016/j.foodcont.2018.03.034.
[33]. Bertolini, F., Ghionda, M. C., D'Alessandro, E., Geraci, C., Chiofalo, V., & Fontanesi, L. (2015). A next generation semiconductor based sequencing approach for the identification of meat species in DNA mixtures. *PloS one*, *10*(4), e0121701. https://doi.org/10.1371/journal.pone.0121701
[34]. Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular ecology*, *21*(8), 2045–2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x
[35]. Ran-Ran Xing, Nan Wang, Ran-Ran Hu, Jiu-Kai Zhang, Jian-Xun Han, Ying Chen, Application of next generation sequencing for species identification in meat and poultry products: A DNA metabarcoding approach, Food Control, Volume 101, 2019, Pages 173-179, ISSN 0956-7135, https://doi.org/10.1016/j.foodcont.2019.02.034.
[36]. Shamsher, M., Knight, A. (2016). Detecting adulterants in food. Food Sci. Technol. Available: http:// www.fstjournal.org/features/30-2/detecting-adulterants-in-food/; Accessed: May 21, 2017.
[37]. Ortola-Vidal, A., Schnerr, H., Rojmyr, M., Lysholm, F., Knight, A. (2007). Quantitative Identification of plant genera in food products using PCR and pyrosequencing technology. Food Control 18:921–927
[38]. Ahmadian, A., Gharizadeh, B., Gustafsson, A.C., Sterky, F., Nyrén, P., Uhlén, M., Lundeberg, J. (2000). Single nucleotide polymorphism analysis by Pyrosequencing. Anal. Biochem. 280:103–110
[39]. Ronaghi, M. (2001). Pyrosequencing sheds light on DNA sequencing. Genome Res. 11:3–11
[40]. Shamsher, M., Knight, A. (2016). Detecting adulterants in food. Food Sci. Technol. Available: http:// www.fstjournal.org/features/30-2/detecting-adulterants-in-food/; Accessed: May 21, 2017.
[41]. Sarri, C., Stamatis, C., Sarafidou, T., Galara, I., Godosopoulos, V., Kolovos, M., Liakou, C., Tastsoglou, S., & Mamuris, Z. (2014). A new set of 16S rRNA universal primers for identification of animal species. Food Control, 43, 35-41