One Year of Next Generation Sequencing (NGS) Data Collection for Food Analysis: Overview of Meat- and Fish-based Samples

Inês Santos[1], Cristina Barbosa[1], Franck Pandiani[1], Mário Gadanho[2], Amanda Manolis[2], Nicole Prentice[2], [1]SGS Molecular, Lisbon, Portugal, [2]Thermo Fisher Scientific, Basingstoke, United Kingdom

ABSTRACT

DNA analysis has gained importance at the service of food authenticity and safety. Next Generation Sequencing (NGS) has been introduced in this sector as a powerful and robust DNA-based method for species identification in food products. NGS requires the development of the correct workflow to ensure results reliability and to make the most of the advantages of this high throughput technique.

For the analysis of meat and fish complex food products we developed a new workflow that uses SGS™ All Species ID kits together with the Ion Torrent™ platform. SGS have successfully implemented this technology into their testing lab and are providing this solution to their clients. In this study data has been collected from a full year of routine analysis. During 2020 a total of 2,100 meat samples and 1,058 fish samples were tested with the mentioned workflow. The samples belong to all kind of food categories, including highly processed products like canned food.

The results show a total of 3,278 identifications for meat species and 1,565 identifications for fish species, being the most common species *Sus scrofa* and *Gadus morhua* respectively. Also, the workflow was applied with an overall successful rate of 99.7% meaning that the DNA extraction was not successful in only 9 samples. This shows that the referred workflow can be implemented for routine analysis contributing to food authenticity and safety.

INTRODUCTION

Next Generation Sequencing testing is a solution based on DNA amplification and sequencing that allows the identification of species present in food and feed samples. This is a critical step to identify food fraud and authenticity verification, improving traceability and food safety at the same time.

SGS Molecular is a pioneer in the introduction of NGS in the food sector through the Ion Torrent™ platform and DNA sequencing apparatus – Ion Chef, Ion PGM and Ion GeneStudio™ S5. With the SGS developed NGS tests it is possible to identify thousands of organisms in complex and processed products. Such a solution relies on the use of combined gene markers stated on the literature, such as Barcode of Life Data Systems (BOLD), to be the most adequate for species resolution.

An NGS workflow (the Thermo Scientific™ NGS Food Authenticity Workflow) was designed to ensure the most universality possible within a target group (meat, plant, fish, bacteria, insects, etc). The method is intended to work as a powerful screening tool for any meat, fish or plant organisms present in a sample. Species-level resolution, or at least the corresponding genus taxonomic level - provides a deconstructed view of a complex product. Additionally, specific strategies are applied to overcome the most common problems associated with food testing: the presence of inhibitory substances and high DNA fragmentation in the samples. The aim is to achieve the most reliable identification results from all kind of matrices.

Taking advantage of the non-targeted and massive sequencing output obtained by NGS, a workflow was developed and tested to identify meat and fish species in food products.

In this study we present the data collected from a full year of analysis in the SGS laboratories showing that the developed workflow can be implemented for routine analysis contributing for food authenticity and safety.

Figure 1. Overview of the complete workflow applied in this study

MATERIALS AND METHODS

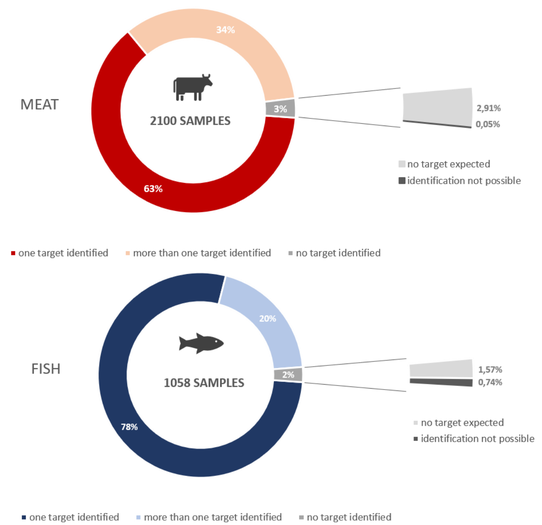
Samples were homogenized and DNA extraction was performed using a commercial kit, NucleoSpin® Food kit (Macherey-Nagel).

The extracted DNA was amplified with the SGS™ All Species ID Meat DNA Analyser Kit or with SGS™ All Species ID Fish DNA Analyser Kit following the instructions. The PCR products were mixed in equal amounts to create the DNA library that was purified with AgentCourt® AMPure® XP beads (Beckman Coulter) according to the manufacturer instructions. The final libraries were quantified with dsDNA BR Assay Kit using the Invitrogen™ Qubit™ Fluorometer equipment (Thermo Fisher Scientific) and sequenced with Ion Chef™ Food Protection Instrument and Ion PGM™ System (Thermo Fisher Scientific) following the instructions.

The amount of DNA sequences generated by the DNA sequencer was very high (between some hundreds of thousands and millions of sequences). Therefore, the data analysis was performed with an internally developed software which contains a set of algorithms that will group the sequences by similarity and compare them with an internal DNA sequences database.

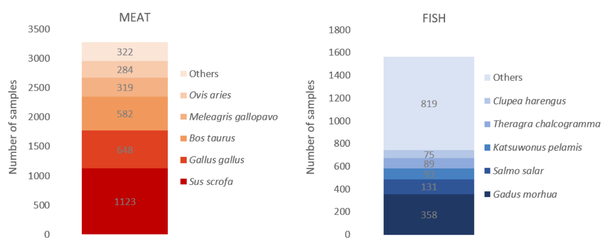
RESULTS

Figure 2. Meat and Fish samples analyzed in 2020 using SGS NGS-based workflow



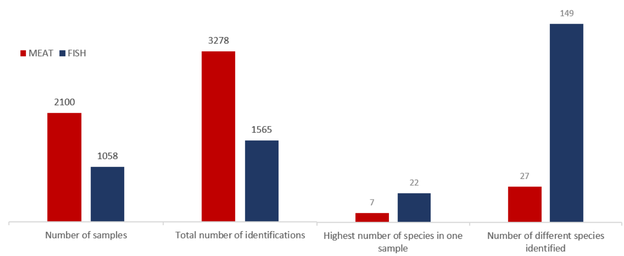
Throughout 2020, 2,100 samples were analyzed for meat targets. In a majority of those samples, only one species was detected, with more than one target identified for a third of the samples. In 64 samples, corresponding to 3% overall of all samples in this analysis, no target was detected, though out of those, 63 were not expected to have any meat targets and only 1 sample lacked identification due to the high degree of processing it had undergone which compromised DNA integrity and further analysis.
Regarding the 1,058 samples processed for fish target identification, the number of those with only one hit was even higher. Once more there was a small percentage for which no target was identified, corresponding to 25 samples in total, out of which only 8 were due to an impossibility to recover analyzable DNA, which can be attributed to the high degree of processing of the samples or the presence of inhibitors.

Figure 3. Top five meat and fish species identified across all samples



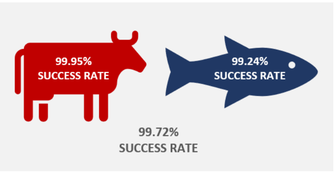
The five main species identified across all samples tested for meat targets, accounting for 90% of all identifications were *Sus scrofa* (wild boar), *Gallus gallus* (red junglefowl), *Bos taurus* (cattle), *Meleagris gallopavo* (wild turkey) and *Ovis aries* (sheep).
For fish, there is a wider diversity of species found, with the five main species identified accounting for only 52% of all identifications. These pertained to *Gadus morhua* (codfish), *Salmo salar* (salmon), *Katsuwonus pelamis* (skipjack tuna), *Theragra chalcogramma* (Alaska pollock) and *Clupea harengus* (Atlantic herring).

Figure 4. 2020 NGS overall data



Of all 2,100 and 1,058 samples analyzed for meat and fish targets, a total of 3,278 and 1,565 identifications were obtained, respectively, accounting for the samples in which more than one species was identified. Overall, the majority of samples only allowed the identification of one species. Samples undergoing fish target analysis were more complex than those for meat identification, with the highest number of species identified in a single sample being 7 for meat and 22 for fish. This can also be observed by the total number of different species across all samples, where 27 different species were identified for meat, with *Sus scrofa* as the most common, and 149 for fish, with *Gadus morhua* as the most common.

Figure 5. NGS-based workflow for the analysis of meat and fish samples



In both analysis a small percentage of samples was impossible to obtain identification for, due to their high level of processing. It is important to note that these correspond only to a low number of complex samples processed, with an overall success rate of 99.7% for this workflow.

CONCLUSIONS

NGS is a promising tool for authenticating many meat and fish samples because:
(1) it is suitable for samples containing highly processed and degraded DNA,
(2) there is no need of a priori species information,
(3) it is cost-effective when processing numerous samples, and
(4) it is possible to detect viable DNA in very low amounts.

We have shown that NGS can be successfully used in complex food matrixes containing meat and fish.

REFERENCES

Goodwin, S., McPherson, J. & McCombie, W. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet 17, 333–351 (2016). https://doi.org/10.1038/nrg.2016.49

Slatko, Barton E et al. “Overview of Next-Generation Sequencing Technologies.” Current protocols in molecular biology vol. 122,1 (2018): e59. doi:10.1002/cpmb.59

Dunisławska, Aleksandra & Łachmańska, Jagoda & Slawinska, Anna & Siwek, Maria. (2017). Next generation sequencing in animal science - A review. Animal Science Papers and Reports. 35. 205-224.

Food Authentication and Traceability. Edit by Charis Galanakis (2020) Academic Press

TRADEMARKS/LICENSING

IAFP, July 2021 LT2667A