Next Generation Sequencing to identify GMO in food and feed products

Sander Willems1*, Marie-Alice Fraiture1,2*, Sigrid C. J. De Keersmaecker1 and Nancy H Roosens1
1 Platform Biotechnology and molecular Biology, Scientific Institute of Public Health, rue J. Wytsmanstraat 14, 1050 Brussels, Belgium
2 Instituut voor Landbouw- en Visserijonderzoek (ILVO), Eenheid Technologie & Voeding (T&V), Burg. Van Gansberghelaan 115, 9820 Merelbeke, Belgium
*equal contribution

The growing number and diversity of GMOs present on the market makes the use of the current for GMO analysis gold standard real-time PCR technology more and more complex and time-consuming in order to detect a GMO in food and feed samples. Indeed, an increased number of screening and event specific methods (targeting the GM element(s) and the junction between the GM cassette and the host genome, respectively) has to be developed and used by the enforcement laboratories in order to cover all the authorised GMOs in a country. In addition, the real–time PCR strategy implies the prior knowledge of the sequence, at least partial, of the GM cassette. Collecting these sequences for unauthorized GMOs is challenging and designing each corresponding method is extremely time-demanding even impossible regarding the numerous possible GM elements found in unauthorized GMOs. This poses a major problem as GMOs remain undetectable when no method targeting the GM element has been used in the tested sample. Recently, to take up the challenge of GMO detection in food and feed matrices, Next Generation Sequencing (NGS) allowing massive parallel DNA fragment sequencing resulting in millions of sequencing reads, has been proposed as a promising technology.

Figure 1: Massive DNA fragments (reads) from a GMO, produced with NGS technology.
In this context, the Scientific Institute of Public Health (WIV-ISP) has started a pioneer study financed by the project UGMMONITOR (SPF convention RF 11/6242) and EPIGMO (Ylieff). This study has first used the NGS technology to assess its potential applicability to detect and identify a GMO (figure 1) in different types of alimentary matrices, i.e. grains of 100% GM rice, grains of 10% GM rice and 100% GM rice noodles (processed food) (figure 2). Secondly, to evaluate the potential use of NGS by “bioinformatics laymen” the data were analysed by using two different platforms: an easy-to-use commercial platform (CLC Genomics Workbench), without the need for a thorough bioinformatics background, and an “in-house” platform allowing greater control of the workflow and parameters and as a consequence demanding a higher level of expertise in bioinformatics. Thirdly, a conceptual statistical framework was developed and applied to estimate the amount of reads necessary to be able to detect and identify several common GMOs at concentrations representative of “typical” food and feed matrices.

Figure 2: Types of rice matrices used in this study: rice grains (a) and “home-made” rice noodles produced from rice grains (b).

Our results showed that it is possible to use NGS to identify and characterize all the types of samples envisaged in this study. The analysis requires only prior knowledge of the sequence, at least partial, of the GM cassette and of the host genome used as reference during mapping of the sequencing reads. Therefore, the NGS strategy allows to use a standardized approach for any type of GMO, in contrast with the specific method development and use that need to be designed for each GMO individually when using the real-time PCR approach. Moreover, a processed matrix such as rice noodles, yielding degraded DNA after DNA extraction from the sample, is not an issue in using the NGS technology (Illumina).

The study highlights also that the development of new user-friendly specific visualization software is necessary to efficiently analyse and deliver the knowledge to the user, especially in the present context of lack of bioinformatics expertise in the enforcement laboratories.
The conceptual statistical model has indicated that a large genome like that of wheat requires a higher number of sequence reads (i.e. a higher coverage or larger sequencing depth), resulting in larger costs per sample but at a price affordable for an "enforcement lab" when the GMO is present at 100%. However, detecting small amounts of GM DNA (1%) in a plant DNA mixture is at the present time impossible when considering the cost and the complexity of the analysis.

The present study offers preliminary information about some major strengths and weaknesses of the NGS technology that need to be addressed before consideration of any routine use of NGS in GMO analysis. It is concluded that NGS has the potential of solving current problems in the GMO detection. However, before any implementation in routine, extended research projects and validation guidelines are necessary.

Detailed results of this research were submitted for peer-reviewed publication.

Acknowledgments

This research is funded by the Federal Public Service Health, Food Chain Safety and Environment (convention RF 11/6242) through the Project UGMMONITOR. The authors would like also to thank Emmanuel Guiderdoni (CIRAD, UMR AGAP, Biological Systems department, Montpellier, France) for his kindness to provide rice grains.

nancy.roosens@wiv-isp.be