European Commission

JRC SCIENCE AND POLICY REPORTS

European Network of GMO Laboratories

Working Group for identification of stacked GM events (WG-IGSE) Working Group Report

2014

Report EUR 26705 EN

Joint Research Centre

EURL European Union Reference Laboratory for GM Food & Feed

European Commission
Joint Research Centre
Institute for Health and Consumer Protection

Contact information
Molecular Biology and Genomics Unit
Address: Joint Research Centre, Via Enrico Fermi 2749, TP 201, 21027 Ispra (VA), Italy
E-mail: eurl-gmff@jrc.ec.europa.eu
Tel.: +39 0332 78 5165
Fax: +39 0332 78 9333

https://ec.europa.eu/jrc

JRC90782

EUR 26705 EN

ISBN 978-92-79-38869-9 (PDF)

ISSN 1831-9424 (print)

doi: 10.2788/87226

Luxembourg: Publications Office of the European Union, 2014

Abstract

This report represents an overview of the current state of the art as discussed during the ENGL Steering Committee meeting of March 2014. The Working Group was asked to remain active and to monitor developments and advances in the field. The European Commission, however, would appreciate a more active role and invites the ENGL to continue its reflection and discussions and to provide regular updates on the progress made in this field. The current report will be reviewed and updated accordingly.

EUROPEAN COMMISSION
JOINT RESEARCH CENTRE

Institute for Health and Consumer Protection
Molecular Biology and Genomics Unit

ENGL-WG for identification of stacked GM events (WG-IGSE)

Working Group Report

The working group

The working group has been chaired by Rupert Hochegger, Austrian Agency for Health and Food Safety, Vienna, Austria. The other members of the working group were: Alexandre Angers, Joint Research Centre (JRC); Gilbert Berben, CRA Walloon Agricultural Research Centre, Gembloux, Belgium; David Dobnik, National Institute of Biology, Ljubljana, Slovenia; Patrick Guertler, Bavarian Health and Food Safety Authority, Oberschleißheim, Germany; Peter Hübert, Landeslabor Schleswig-Holstein, Neumünster, Germany; Joachim Kreysa, Joint Research Centre (JRC); Antoon Lievens, Joint Research Centre (JRC); Marco Mazzara, Joint Research Centre (JRC); Jaroslava Ovesna, Crop Research Institute Reference Laboratory for GMO Detection, Prague, Czech Republic; Alex Patak, Joint Research Centre (JRC); Sven Pecoraro, Bavarian Health and Food Safety Authority, Oberschleißheim, Germany; Mauro Petrillo, Joint Research Centre (JRC) and Maddalena Querci, Joint Research Centre (JRC).

Foreword

This report represents an overview of the current state of the art as discussed during the ENGL Steering Committee meeting of March 2014. The Working Group was asked to remain active and to monitor developments and advances in the field. The European Commission, however, would appreciate a more active role and invites the ENGL to continue its reflection and discussions and to provide regular updates on the progress made in this field. The current report will be reviewed and updated accordingly.

Executive summary

The European Network of GMO Laboratories (ENGL) was requested to provide technical advice on how to improve the differentiation between stacked and single events, and a Working Group "identification of stacked GM events" (WG-IGSE) was formed to review the current state of play, explore the feasibility of novel approaches under routine analytical conditions and propose research strategies.

The WG addressed the term "stacked GM events" to GMOs containing more than one GM event combined via conventional crossing of previously existing GMOs. A literature search was performed to review prior research on the problem of stacked GM events detection. The only existing strategies outlined in the literature able to differentiate between the stacked GM events and mixtures of their parental events were by testing samples that were known to contain only genomic DNA extracted from a single organism, for example a single maize kernel.

During its discussions, the WG identified different potential approaches that could be explored in order to differentiate between stacks and combinations of single GM events. The identified approaches were grouped conceptually. The conceptual groups were marker-assisted identification, single-cell analyses, and statistical approaches; the pros, cons and feasibility of each of these approaches were discussed.

Single-cell analyses approaches were identified by the WG as potentially able to unequivocally prove the presence of stacked events in cases where intact cells/nuclei are present in the sample. However, they would also be the ones that require the most research and development to adapt and standardise the protocols as well as to test their feasibility. Marker-assisted identification, supported by Next Generation Sequencing, was also identified as promising. Again, there is still a large amount of work required to adapt these techniques and to develop efficient bioinformatics pipelines. If successfully identified and selected, these markers could be screened using conventional methods. Statistical approaches could be applied to the analyses as they are currently performed; however, they cannot in all cases directly detect stacked events in the sense of irrefutable proof of their presence in a sample and only provide an indication on how well the presence of a stacked GMO explains the observed results.

Because of the large amount of preliminary work required, all of the approaches offer limited short-term perspectives for routine and large scale implementation. As a consequence, the WG recommends that, in the short term, these proposals are used for case-by-case studies in research environments to establish which of the proposed approaches, if any, are suitable for the problem addressed by the WG: the differentiation of stacked GM events from mixes of their parent events. It would also contribute to the development of the required specific protocols. Stack identification could be approached in a step-by-step manner, starting with statistical approaches and the subsequent steps determined by the strength of stack presence indication and the nature of the sample.

Contents

1 Introduction

1.1 Detection of stacked GM events – problem formulation

Legislations in the European Union (EU) foresee stringent requirements for GMO approval, labelling and traceability. The availability of analytical methods that allow for a sensitive and accurate determination of GMO content is then a key element in the implementation of these requirements.

The polymerase chain reaction (PCR) has proven to be the most accurate technique available and to accommodate samples ranging from seeds to highly processed food/feed. For this reason the current detection methods used by the European Network of GMO Laboratories (ENGL) members are based on this technique. Typical targets include common elements in GMO constructs such as the CaMV 35S promoter (P-35S) and the terminator of the *nopaline synthase* gene of *Agrobacterium tumefaciens* (T-nos). However, since these elements are shared between different GMOs, the highest levels of specificity is achieved by event-specific methods that target the junction between the inserted DNA and the recipient genome, as this region is unique to each DNA integration event.

As the number of introduced dominant GM traits, comprised mainly of herbicides tolerance and pests resistance, continued to increase within the same crops, products were generated that combine more than one of these traits in the same organism. This process is called "gene stacking" and can occur in different ways (reviewed in Taverniers *et al*., 2008). When multi-trait events are obtained by conventional crossing of individual parental GMOs, the resulting GMO contains a combination of foreign DNA insertion sites carried by the chromosomes of the parental GMOs. For this reason, the sequences at the border of the inserts and the surrounding host genome, which are the target of the parental lines event-specific methods, will be found unchanged in the resulting multi-trait GM event and will thus be detected when testing DNA extracted from the stack with the event-specific methods.

As a consequence, a parental event-specific method will then produce a positive signal with all stacked GMOs that result from the crossing of this event with other GMOs (or other stacks that resulted from such a previous crossing). The same applies to all element- or construct-specific methods that would detect the parental event. Stacked GM events presence can be tested for through the combination of the parental event-specific methods. However, this method only allows for the discrimination between a sample containing a stacked GMO AxBxC, and a sample that contains a mixture of the single events A, B and C or even a mix of these single events and the stack in cases when single seeds or cells can be tested.

For GMOs containing more than one GM event obtained in ways other than conventional crossing, for example by co-transformation or re-transformation, integration of the inserts ("cassettes") into different loci of the host genome compared to the parental lines would occur. This makes it then possible to develop detection methods specific for these novel insertion sites, thus to directly identify the resulting multi-trait GMO with a specific method, targeting the junction of the insert and the loci. This is the case even when the same transformation vector previously used for existing GMOs is used in the co- or re-transformation.

An example for this last point is the GM cotton MON 15985, which was produced by re-transformation of GMO event MON 531 in order to add the expression cassettes for the Cry2Ab2 and GUS proteins. A detection method was developed and validated that is specific for this multi-event GMO (QT-EVE-GH-005 in the JRC Methods database) by targeting the insert junction of the new insertion site.

However, it should be kept in mind that these cases are not completely immune to the issues of stacked and single events discrimination. For example, possible improvements in targeted homologous recombination for transgenic cassettes could eventually restrict the randomness of the genomic insertion site, although whether this could be achieved in a way that would affect the specificity of existing event-specific detection methods remains to be seen in practice. More concretely, there exists the possibility that additional insertion sites generated by re-transformation could be segregated and isolated from the original insertion sites during conventional breeding. To return to the example of MON 15985, it would be possible to segregate the additional insertion site (called MON 15947 by Monsanto) into a cotton plant that no longer contains the original MON 531 insert. In that case, the QT-EVE-GH-005 detection method would recognize both MON 15985 and the new plant that only contains the insert it targets, and the same problems addressed by this report would be faced.

1.2 Mandate of the WG

To address this issue, the decision to form a Working Group (WG) for "identification of stacked GM events" was taken during the 25th ENGL Steering Committee meeting (held on the 11th of September, 2013). This WG was given the task to draft a report by March 2014 for the 26th ENGL Steering Committee.

The mandate of the WG was drafted, tasking the WG with:

- reviewing the current state of play (e.g. by literature search)
- exploring the feasibility of novel approaches under routine analytical conditions
- proposing research strategies
- providing recommendation(s)

1.3 Working definition of stacked GM events

In order to properly frame its task, the WG first discussed a working definition of the "stacked GM event" for which they felt necessary to explore the feasibility of novel approaches to discriminate these organisms from mixtures of parental events. For the reasons explained in section 1.1, it was agreed that for the purpose of this WG the following working definition should be used:

"*A stacked genetically modified (GM) organism is defined as a GMO containing more than one GM event combined via conventional crossing of previously existing GMOs*".

This working definition is consistent with the terminology described in Taverniers *et al*. (2008), that labels "stacked events" as the product of conventional crossing between existing independent events, and as such, a special subset of "multiple events" (i.e. a GMO carrying multiple modifications, at one locus or different loci).

Although this definition guided the discussions for the approaches described in the report, these approaches should apply to the other concerns described in section 1.1 such as re-transformation or co-transformation if, for the latter two techniques, the so inserted new traits might be found independently as segregants in commercialized products.

2 Literature review

A literature search was performed to review prior research on the problem of stacked events detection, spanning both the Scopus[1] and Pubmed[2] databases. The following articles were identified:

- Akiyama, Hiroshi, et al. "Quantitative detection system for maize sample containing combined-trait genetically modified maize." Analytical chemistry 77.22 (2005): 7421-7428.
- Wu, Y et al. "Event-specific qualitative and quantitative PCR detection methods for transgenic rapeseed hybrids MS1xRF1 and MS1xRF2" J. Agric. Food Chem. 55 (2007): 8380-8389.
- Xu et al. "Event-specific detection of stacked genetically modified maize Bt11xGA21 by UP-M-PCR and real-time PCR" J. Agric. Food Chem. 57 (2009): 395-402.
- Choi, Sun Hee, et al. "Development of detection system using multiplex PCR and liquid beadarray for stacked genetically modified rice event (LS28× Cry1Ac)." Journal of the Korean Society for Applied Biological Chemistry 53.5 (2010): 639-646.
- Kim, Su-Youn, et al. "Detection system of stacked genetically modified maize using multiplex PCR." Food Science and Biotechnology 19.4 (2010): 1029-1033.
- Choi "Hexaplex PCR assay and liquid bead array for detection of stacked genetically modified cotton event 281-24-236x3006-210-23" Anal Bioanal Chem. 401 (2011): 647-655.
- Akiyama, Hiroshi et al. "Quantification and identification of genetically modified maize events in non-identity preserved maize samples in 2009 using an individual kernel detection system" Food Hyg. Saf. Sci. 53.4 (2012): 157-165.
- Shin, Kong-Sik, et al. "Event-specific detection system of stacked genetically modified maize by using the multiplex-PCR technique." Food Science and Biotechnology 22.6 (2013): 1763-1772.

[1] http://www.scopus.com/

[2] http://www.ncbi.nlm.nih.gov/pubmed/

- Querci, Maddalena, et al. "Real-time PCR-based ready-to-use multi-target analytical system for GMO detection." Food Analytical Methods 2.4 (2009): 325-336.

Whether or not it is explicitly mentioned in the title, most of these articles involve the use of the multiplex PCR technology, i.e. the combination of multiple pairs of primers in the same PCR reaction. The only exception is Wu *et al.* (2007), who run the multiple PCR in parallel reactions.

Differences are also observed in the way the amplification results are analysed: agarose gel electrophoresis (Xu *et al.,* 2009 (for screening); Kim *et al.,* 2010, Shin *et al.,* 2012), TaqMan™ probes (Akiyama *et al.,* 2005; Wu *et al.,* 2007; Xu *et al.,* 2009 [for quantification]), liquid bead arrays (Choi *et al.,* 2010; Choi, 2001) and microchip electrophoresis (Akiyama *et al.*, 2012).

In all the articles, the same general strategy is used for the generation of the different primers, which is to produce a panel of primer sets that would recognise elements from all the different stacked insertion sites. Positive results with all the primers sets would imply that all the insertion sites are present in the sample, hence suggesting a stacked GMO.

However, as described above, this does not solve the challenge of differentiating between a sample containing a stacked GMO and a sample containing a mixture of the corresponding parental single GM events. The only way to differentiate between the two cases with these strategies would be to start with a sample that is known to contain only genomic DNA extracted from a single organism, for example a single maize kernel. This strategy is described in Akiyama *et al.* (2005), Akiyama *et al*. (2012) and Shin *et al*. (2012). Choi (2011) mentions in the conclusions that the technique described could be applied "for differentiating stacked events from mixed samples containing several single-trait events", but the way this could be achieved is unclear and not specified in the article.

The only exception is the method described in Xu *et al.* (2009), that uses mathematical equations and the combination of results obtained from multiple qPCR analyses to identify and quantify the amount of stacked GMOs and the corresponding individual events in a mixed sample, in this case maize Bt11, maize GA21 and the Bt11xGA21 stacked GMO. The technique depends on the fact that the inserted cassettes of these two single events contain different copy numbers of the same element, i.e. the nopaline synthase terminator (T-nos): the Bt11 insert contains 2 copies, the GA21 insert 1 copy, and thus the Bt11xGA21 stack 3 copies. Comparing the copy number of T-nos in the sample to the copy number of an endogenous gene allows, through computations using the described mathematical formulas, to infer the relative amount of each organism in the original sample.

In addition, the authors explain an additional limitation of the technique, which is that it can only be applied when the mixed sample is known to contain only three of the following four organisms: Bt11, GA21, Bt11xGA21 and non-GM maize. The exact organisms that are present need to be known in advance since the formulas vary depending on the type of mixture.

In summary, although a progress from single-kernel detection, the described method cannot be generalised.

3 Potential approaches

During its discussions, the WG identified different potential approaches that should be explored with regard to their capacity to differentiate between stacks and combinations of the single GM events combined in the stack. The identified approaches were grouped conceptually and are briefly defined here. Discussions on the pros, cons, and feasibility of each of these approaches are found in section 4, below.

3.1 Marker-assisted identification

At the phenotype level and using the developed event-specific detection methods, stacks are indistinguishable from the sum of the single events.

One possibility then involves the identification (and subsequent detection) of specific differences ('markers') at the metabolic or molecular level between stacks and their parental events, in addition to the phenotypes and/or outside the currently targeted genomic regions.

The various types of differences that can be explored are described below:

3.1.1 Metabolic markers

Independently from the GMO-specific phenotypic traits, the additional breeding process required to produce stacked events may create an organism with a different final metabolic profile from the original parental events. These differences are not a direct consequence of the genetic modification, and are known to exist between different varieties of the same species. Once characterised, these metabolic patterns could be used to differentiate stacked events and mixtures of the parental organisms.

3.1.2 Plant genomic markers

By inbred line production, the genetic background of the organism is modified, changing the patterns of Nucleotide Polymorphism, Single Nucleotide Polymorphism (SNPs) and Insertion/Deletions (In/Dels), in the plant genome. Despite the fact that stacks and events share the transgenic insert, this leads to a characteristic pattern in the rest of the genome that can be monitored and detected.

3.1.3 Patterns of secondary insertion sites

Some parental events contain secondary insertion sites of the targeting cassette in their genome in addition to the main insert. These secondary insertion sites are not selected during the crossings that are made to produce the stacks. It is then possible that the pattern of these secondary insertion sites is different between stacks and their corresponding parental events. Developing methods that specifically detect specific patterns of secondary insertion sites could then differentiate stacks and mixtures of single events.

3.1.4 Insert sequence polymorphisms

Like any other part of the genome, the DNA sequence of the transgenic insert is subject to random mutations that can accumulate during the production of recombinant inbred lines. Comparisons of the sequences between stacks and parental events may identify these differences (SNPs and In/Dels), which can then become markers for differentiating stacks and single events.

3.2 Single-cell analyses

As described in the literature review, one way to differentiate stacks and a mix of parental events is to start the analyses from a single kernel, as any event identified in this case is guaranteed to be stacked in this organism. In practice, though, this situation is fairly rare as samples to be analysed are often processed to some degree. However, should the processed sample allow for the isolation of intact cells (or nuclei), the same concept could be applied with the appropriate techniques.

3.2.1 Fluorescence in situ hybridisation (FiSH)

The "stacked" nature of the event can be confirmed by looking directly at the inserts found in the nuclei of the organism, using a technique called fluorescence *in-situ* hybridisation (FiSH). FiSH is a technique that was specifically developed to detect the presence or absence of a specific DNA sequence on the chromosomes inside nuclei by directly hybridising a fluorescently-labelled DNA probe to the denatured DNA in the fixed nuclei.

Through this technique, it might be possible to distinguish between stacked GM events and single parental GM events by generating probes specific for each of the parental events transgenic regions and capturing images of the stained nuclei.

3.2.2 Single-cell PCR

Using the existing event-specific PCR detection methods on single cells, the analysis becomes straightforward since a mixture of positive results unambiguously show that the analysed sample contains a stack of all the detected genetic modifications.

3.3 Statistical approaches

3.3.1 Relative signal analyses

The proposed method is comprised of a series of calculations based on the qPCR results of the quantification of the separate events and determines how well the presence of a stacked event is capable of explaining the observations. The result of this approach is a measure of how well the hypothetical presence of a stacked event fits the observed quantification data.

3.3.2 Digital PCR distribution assessment

The approach involves using digital PCR (dPCR) to determine the segregation patterns of the individual targets. In case of a stacked GM event where both constructs are close enough to each

other on the same chromosome, the linkage between the single GM events might be evidenced in digital PCR.

3.3.3 Subsampling distribution assessment (Seedcalcstack)

Segregation of the positive GM events signals can also be tested at the level of subsamples. Seedcalcstack, an application developed by ISTA for seed lot purity testing, contains an implementation of this concept (using the maximum likelihood approach) with up to three stacked events in a conventional seed lot[3].

4 Feasibility and applicability

Each of the approaches described in section 3 has its own set of pros and cons, described below.

4.1 Marker-assisted identification

A general pro of these approaches is the fact that, once differences between stacked GMOs and parental events are identified, standardised and well-established methods exist to target them. However, a general con of these approaches is that it relies on the assumption that these differences do exist, are stable, and can be efficiently detected for all combinations of stacks and their parental events. It also requires the initial characterisation of the GM organisms in order to identify these differences, and this work can be labour intensive.

Various sensitive and specific assays exist to identify and quantify metabolic markers; however, these involve techniques such as MALDI-TOF and other mass spectrometry, as well as other proteomic/metabolomics techniques. These require specific equipment that is not currently part of the GMO detection systems. Guidelines have been developed to characterise seeds (outside the context of GMO detection and quantification, see ISTA guidelines[4]) that could be used to develop techniques that can be applied to single events and stacked organisms. Once established, such a method could also allow quantification of the stacks and single events mixed in the same sample.

Plant breeders have been performing testing and selection of specific nucleotide polymorphisms in plant organisms using Next-Generation Sequencing (NGS) for a couple of years. Guidelines have been developed for seed approval[5] and within the ISTA working programme[6]. A timeline of the different molecular markers targeted in these studies is shown in the figure below, taken from Henry (2012).

[3] http://www.seedtest.org/en/statistical-tools-for-seed-testing-_content---1--1143--279.html

[4] http://www.seedtest.org/en/home.html

[5] http://www.upov.int/test_guidelines/en/list.jsp

[6] http://seedtest.org/upload/cms/user/ECOM11011ApprovedWorkingProgramme2010-2013_VAR.pdf

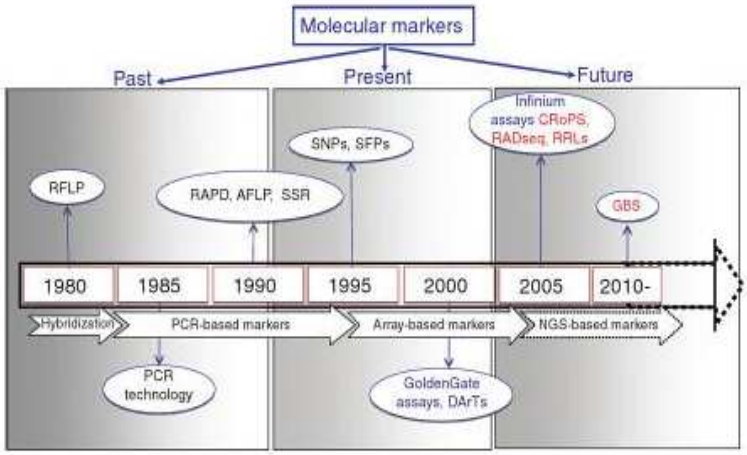

Figure 10.1 Paradigm shift in marker discovery: from hybridization-based RFLPs to NGS-based high-throughput markers. Markers have been classified into past, present, and future molecular markers. Markers highlighted with blue color are micro-array-based markers, while as those highlighted with red color are NGS-based markers.

For maize, soybean, wheat and rice, efforts are also currently underway to characterise varieties based on microsatellites (SSR)[7], or allele specific (e.g. KASP, developed by LGC genomics) markers, techniques that could also be used in the scope of stacked GM events detection to identify the detectable markers.

A good knowledge of the variety-specific genome sequences is a prerequisite for efficient polymorphisms identification and these methods require the handling of large amounts of information. The different currently used techniques, each with their own pros and cons, include linkage disequilibrium (LD) and nested association mapping (NAM) (Sajjad et al., 2014). A large amount of evidence and experience can be found for the use of these markers for breeding purposes, and it remains to be verified whether they can be applied for the identification (and differentiation) of stacked and parental events. In contrast to bi-parental linkage studies, LD based association studies can identify variants with relatively small individual contributions. Association mapping (AM) is currently the method of choice to exploit variations found in cultivars from multiple as well as regional breeding programs. As detailed measurement of genetic variations is required, this problem could only be solved by Genome Wide Association Studies (GWAS) using NGS. However, extensive preliminary work is still required to determine how to characterize the traits required for stacked event identification, collect the analytic tools for evaluation, interpret the data and harmonize the methodology in the EU.

Once the differences are identified, it is not yet clear how many of these markers need to be targeted for the specific need of GMO identification. As a current order of magnitude, about 1100 markers are sufficient to determine about 5000 recombinant inbred lines, using NAM, see Yu *et al.* (2008). It is possible that the number of markers will be suitable to be analysed using qPCR, the main currently used technique for GMO identification and quantification. Otherwise, detection could be

[7] https://www.seedtest.org/upload/cms/user/OGM13-03ActivityReportoftheISTACommittees2012_VAR.pdf

achieved using microarrays and/or NGS, which would increase the complexity of large-scale implementation of this technique. For additional details about large-scale analyses, see a recent review (Bohra, 2013).

The first two approaches (metabolic markers and genomic SNPs) involve the independent characterisation of single events and stacked GMOs as unique organisms (parental linkage approximation), standalone of their genetic modifications. These analyses should also take into account the fact that usually the same GMO events exist in many varieties of the same crop.

The other two approaches (patterns of secondary insertion sites and insert sequence polymorphisms) specifically target the genetic modifications of the GMOs.

Some existing events have been shown to contain secondary inserts besides the main insertion site that is targeted by the detection method. During the breeding to produce an organism with stacked primary insertion sites, it is possible that some of these additional insertion sites were also taken by the stacked GMO. Since these secondary insertion sites have unique integration sites, it is possible to develop specific qPCR methods to detect and quantify them, as is currently done for the main inserts. However, this technique would require that 1) the parental lines have secondary insertion sites that are molecularly characterised and 2) the stacked events have separate and unique patterns of presence of these additional sites. In addition, since these sites can be lost, but not gained, during the breeding to produce the stacked organism, detection of stacks can only be done through analyses of “negative” detection results.

An advantage of this approach is that a proof of concept experiment can be designed with the currently available information. Two events for which secondary insertion sites are reported and characterised, GM event 40-3-2 Roundup Ready soybean (MON-04032-6) and GM event 305423 soybean (DP-305423-1), are also found together in a stack (305423x40-3-2 soybean). Primers could be designed to detect the insert region of the secondary insertion sites to determine the pattern of presence of these sites in each of the organisms. If differences are observed, they could be used for determining, in an unknown sample, the content of parent and stack organisms, respectively.

In practice, the applicability of this method is limited by the mandatory presence of characterised secondary insertion sites in the parental GM events.

The theory and limitations of the “Insert sequence polymorphism” are similar to the plant genome polymorphisms. Since the sequence analysed is much smaller (insert vs whole genome), the work required to identify, characterise and detect the polymorphisms is simpler. For the same reason, however, the odds of identifying such a difference are considerably smaller.

4.2 Single-cell analyses

A general pro of these approaches is the fact that, if successful, they provide a direct demonstration of the existence/presence of the stacked organism in the sample. However, due to the requirement of isolating intact nuclei, they cannot be applied to samples that underwent some types of processing (e.g. cooking or boiling). In addition, although made at the level of the individual cells,

those analyses must be adapted for a sufficient throughput such that a significant number of cells are tested, in order to correctly analyse samples not consisting in a pure organism.

For FiSH, protocols for staining plant cells already exist in the literature. Furthermore, FiSH has already been used to localise transgenes in GM plants (Santos *et al.,* 2006; Choi *et al*., 2009; Suzuki *et al.*, 2011). However, these experiments were not developed for routine GMO detection or analysis, and protocols would need to be adapted and standardised, including for the preparation of samples from different matrices (grain, flour, etc.).

Once the protocols are established, probes need to be developed that would specifically detect the inserts of each of the known GM events and this allows the identification of stacked GM events by nuclear co-staining. There are practical limitations in the number of single events (and their associated stacks) that can be simultaneously detected due to the requirement of using a different fluorophore for each probe, also taking in consideration the autofluorescence of the matrix.

In addition to the prerequisite of adapting/developing the protocols, there is a need for specific equipment (i.e. fluorescence microscope and image capture equipment) that is not currently used for GMO detection. High-throughput image captures and analyses, such as those used in High Content Screening (Haney, 2008) could be used to analyse a sufficient number of cells.

For single-cell PCR, one possibility to gather the single cells is by isolating protoplasts from the plant tissue and then sorting them into separate reactions, for example by nuclei sorting (Zhang *et al.,* 2008). The main obstacle here is the low throughput, since one cell represents one reaction.

To overcome this, the reaction could be done via digital droplet PCR (ddPCR). In ddPCR, the PCR mix is randomly distributed into a large number of partitions (> 10.000), which contain either zero, one or more copies of the target sequence. These partitions undergo a PCR and fluorescence is measured individually resulting in positive and negative droplets. For this, the sample of cells or nuclei is added to the PCR reaction mixture and divided in smaller volumes by making an emulsion. However, digital droplet PCR as it currently exists was not designed for single-cell analyses, as intact cells and nuclei are typically larger than the average droplet size. As a consequence, other types of emulsion PCR, using conventional PCR machines, should then be used (as described, for example, by Guo *et al.,* 2011).

Finally, one PCR could be run on a sample composed of multiple nuclei simultaneously and the amplification within each nucleus measured by FiSH or flow cytometry. The problem here would be how to restrict the reactions to within the nuclei. This type of reaction performed inside plant cells has already been described (Woo *et al.,* 1995) but the protocols need to be adapted to each specific type of cells.

4.3 Statistical approaches

The purpose of these approaches is to provide an indication on how well the presence of a stacked GMO (compared to individual parental events) explains the observed results. Thus, it is important to point out that these approaches do not aim at directly detecting stacked events in the sense of

irrefutable proof of their presence in a sample. That said, they would still provide important information that could be used in decision-making for the downstream analyses in GMO identification.

The relative signal analyses approach involves investigating the confidence bounds of the event quantification (that can be obtained in different ways, see below). Overlaps may indicate that two or more events are present in equal copy numbers. If these confidence intervals do not overlap, it is possible to calculate the observed GM amount that could be explained by the presence of a stacked event (e.g. the presence of a stacked organism could explain 90% of the observed GM content in the sample).

The confidence intervals can be based on a parametric bootstrap approach in order to minimize the amount of laboratory work needed. By bootstrapping the regression analysis of the calibration curves and the observations in the sample, confidence intervals for the quantification can be obtained using the bias-corrected and accelerated method (Efron *et al.,* 1987). The data for these analyses could be provided by the currently implemented qPCR-based detection methods, in particular using highly uniform platforms, such as the pre-spotted plates developed by the JRC (Querci *et al.,* 2009).

In addition, digital droplet PCR (ddPCR, see above) can also provide input data for the relative signal analyses approach in the form of absolute copy numbers of GM events in a sample (Burns *et al.,* 2010; Corbisier *et al.,* 2010; Žel *et al.,* 2012; Morisset *et al.,* 2013). These can be calculated directly from the ratio of positive and negative droplets and fed in the analyses using binominal or Poisson distribution.

The main concept of using the digital PCR (dPCR) to differentiate between stacked and mixtures of individual events involves the analyses of the segregation pattern of the positive signals for each of the target sequences. The latter may be more similar in stacked events (i.e. found in the same partitions) compared to mixed single GM events. However, to be detectable, the observed differences must be higher than the bias of the dPCR and the loci of the different insertion sites should be on the same chromosome at less than 50 cMorgan from each other. In addition, this procedure might be highly dependent on the sample preparation (fragmentation).

The current generation of dPCR machines offers a limited multiplexing capability, which limits the number of single events that can be tested simultaneously. In addition, dPCR is so far not available in most control laboratories and is, depending on the type of device, expensive to acquire and use.

A similar principle can be applied without the use of dPCR if intact seeds or kernels are available for testing. Instead of droplets, subsamples of kernels are tested for each of the single GM events. Results are combined for analyses, for example using Seedcalcstack, an R application linked to an Excel template. The composition of single events/stacks in the original sample can then be evaluated with the maximum likelihood approach (Aldrich, 1997). The requirement for intact kernels in the sampling step represents the main limitation of this method, similar to the single-kernel techniques described in the literature review. Despite this, the statistical background and techniques developed

in the Seedcalcstack project could be of interest if adapted and applied to the other approaches described here.

4.4 Summary Table

The pros and cons of the different proposals described in the text are summarised in the following table:

	Pros	Cons	Equipment required	Applicability
3.1.1. Metabolic markers	Available assays are very sensitive and specific. Allows quantification of stacks and single events event if mixed in the same sample.	Not compatible with currently used GMO detection approaches and equipment. Not based on inserted DNA sequence, so relies on the existence and stability of markers across different varieties of the same GM event (cross-breeding).	Mass spectrometer Equipment for other proteomic/ metabolomics techniques	Most likely restricted to raw (intact) material depending on markers stability.
3.1.2. Plant genome markers	Differences are guaranteed to exist. Could be compatible with PCR if the number of markers is sufficiently low.	Extensive work required to obtain genome information for all GM events and stacks and each variety in which they are commercialized.	Next-Generation Sequencer Microarrays	All materials. Might be more amenable to some species depending on the breeding techniques. Restricted to stacked and parental single events for which LD and AM data, generated by GWAS, are available.

3.1.3 Secondary insertion sites	Identical to currently used GMO detection methods.	"Negative" detection since stacks can only lose, but not gain, secondary insertion sites. Would require the characterization of all commercially available varieties of the stacks and their parental single events.	qPCR instrument	All materials. Restricted to stacked events comprised of at least one single parental event with a characterized and stable secondary insertion site.
3.1.4 Insert sequence SNPs	Also allows quantification. Compatible with currently used GMO detection methods.	Differences are not guaranteed to exist.	qPCR instrument Microarrays	All materials. Restricted to single events and stacks for which a polymorphism in the event insert exists and is characterised.
3.2.1 FiSH	Direct, visual confirmation of the stacked organism. Universal for stacked events, only specific probes for single events have to be designed.	There is a limitation in the number of events that can be analysed at the same time. Existing protocols need to be adapted. Low throughput.	Fluorescence microscope and image capture instrument	Material with intact nuclei.

3.2.2 Single cell PCR	Direct indication of stack presence. In case of a digital version quantification could be possible.	Amplification reaction should be done in multiplex (individual cells cannot be divided into separate reactions). Preparation of cells could be laborious and tricky. Low throughput.	Cell sorter, flow cytometer PCR instrument	Material with intact nuclei.
3.3.1. Relative signal analyses	Easy to implement using the results of currently performed testing.	Loci requirements may severely limit applicability. Do not give absolute certainty about the presence or absence of stack. Uncertainty of the technique increases when low amounts of GMOs are in the sample.	PCR instrument	All materials.
3.3.2 Digital PCR	Rapid and is easy to perform, therefore can be applied for routine analyses Qualitative, as well as quantitative Highly sensitive at low copy number concentrations Not as prone to PCR inhibition as qPCR analysis	Does not give any certainty about the presence or absence of stack. Differences regarding the segregation pattern must be higher than the bias of the dPCR Limited multiplexing capability (currently duplex PCR)	dPCR instrument	All materials Currently limited to double stacks

3.3.3 Subsampling distribution assessment (Seedcalcstack)	Provides the most likely composition of single events/stacks in the original sample. Already developed, implemented and used.	Specific sampling requirements	PCR instrument	Limited to grains/seeds/kernels Currently limited to double or triple stacks

5 Conclusions and recommendations

Several approaches to identify stacked GM events, based on different technologies, have been compiled and discussed in this document by the WG. However, these approaches offer limited short-term perspectives for routine and large-scale implementation due the fact that they require significant additional research in order to adapt the techniques for this specific application and to determine their suitability and applicability. In addition, these approaches generally require specific and sometimes expensive equipment not currently used in routine GMO testing.

It should be emphasised that, for most of the approaches, detailed information on events sequence is required. In addition, access to certified reference materials for both single events and stacks is crucial for their implementation and future use, and these are currently not available for most stacked GM events.

The only techniques identified by the WG that can unequivocally prove the presence of stacked events are the approaches described in the single-cell analyses sections. However, they would require significant research and development to adapt and standardise the protocols and to test their feasibility. These approaches are also limited to cases where intact cells/nuclei are present in the sample.

Another promising technique identified is Next Generation Sequencing. This technique produces a large body of sequence information that can be used to identify detectable differences (sequence polymorphisms) between the stacked event and its parental single event organisms. Again, there is still a large amount of work required to adapt these techniques and to develop efficient bioinformatics pipelines. Once identified and selected, some of these markers could be screened using conventional methods.

The WG recommends that, in the short term, these proposals are used for case-by-case studies in research environments. This experience could then be used to establish which of the proposed approaches, if any, are suitable for the problem addressed by the WG: the differentiation of stacked GM events from mixes of their parent events. It would also contribute to the development of the required specific protocols. Following this, stack identification could be approached in a step-by-step manner (workflow), starting with statistical approaches (that can be applied to the information as it is currently obtained) and the subsequent steps determined by the strength of stack presence indication and the nature of the sample.

6 References

Akiyama, Hiroshi et al. "Quantification and identification of genetically modified maize events in non-identity preserved maize samples in 2009 using an individual kernel detection system" Food Hyg. Saf. Sci. 53.4 (2012): 157-165.

Akiyama, Hiroshi, et al. "Quantitative detection system for maize sample containing combined-trait genetically modified maize." Analytical chemistry 77.22 (2005): 7421-7428.

Aldrich, John. "RA Fisher and the making of maximum likelihood 1912-1922." Statistical Science 12.3 (1997): 162-176.

Bohra, Abhishek. "Emerging Paradigms in Genomics-Based Crop Improvement." The Scientific World Journal 2013 (2013).

Burns, M. J., A. M. Burrell, and C. A. Foy. "The applicability of digital PCR for the assessment of detection limits in GMO analysis." European Food Research and Technology 231.3 (2010): 353-362.

Choi, Hae-Woon, et al. "Stability and inheritance of endosperm-specific expression of two transgenes in progeny from crossing independently transformed barley plants." Plant cell reports 28.8 (2009): 1265-1272.

Choi, Sun Hee, et al. "Development of detection system using multiplex PCR and liquid beadarray for stacked genetically modified rice event (LS28× Cry1Ac)." Journal of the Korean Society for Applied Biological Chemistry 53.5 (2010): 639-646.

Choi, Sun Hee "Hexaplex PCR assay and liquid bead array for detection of stacked genetically modified cotton event 281-24-236x3006-210-23" Anal Bioanal Chem. 401 (2011): 647-655.

Corbisier, Philippe, et al. "Absolute quantification of genetically modified MON810 maize (Zea mays L.) by digital polymerase chain reaction." Analytical and bioanalytical chemistry 396.6 (2010): 2143-2150.

Efron, Bradley. "Better bootstrap confidence intervals." Journal of the American Statistical Association 82.397 (1987): 171-185.

Guo, Jinchao, et al. "MPIC: a high-throughput analytical method for multiple DNA targets." Analytical chemistry 83.5 (2011): 1579-1586.

Haney, Steven A., ed. High Content Screening: Science, Techniques and Applications. John Wiley & Sons, 2008.

Kim, Su-Youn, et al. "Detection system of stacked genetically modified maize using multiplex PCR." Food Science and Biotechnology 19.4 (2010): 1029-1033.

Morisset, Dany, et al. "Quantitative analysis of food and feed samples with droplet digital PCR." PloS one 8.5 (2013): e62583.

Querci, Maddalena, et al. "Real-time PCR-based ready-to-use multi-target analytical system for GMO detection." Food Analytical Methods 2.4 (2009): 325-336.

Sajjad, Muhammad, Sultan Habibullah Khan, and Rashid Mehmood Rana. "Family and/or Friends? Gene Mapping at Crossroads." American Journal of Plant Sciences 5 (2014): 112.

Santos, Ana P., et al. "In situ methods to localize transgenes and transcripts in interphase nuclei: a tool for transgenic plant research." Plant Methods 2.1 (2006): 18.

Shin, Kong-Sik, et al. "Event-specific detection system of stacked genetically modified maize by using the multiplex-PCR technique." Food Science and Biotechnology 22.6 (2013): 1763-1772.

Suzuki, Go, et al. “Transgenic rice plants harboring the grain hardness-locus region of Aegilops tauschii.” Plant Cell Reports 30.12 (2011): 2293-301.

Taverniers, Isabel, et al. "Gene stacking in transgenic plants: towards compliance between definitions, terminology, and detection within the EU regulatory framework." Environmental biosafety research 7.04 (2008): 197-218.

Woo, Ho-Hyung, Lindy A. Brigham, and Martha C. Hawes. "In-cell RT-PCR in a single, detached plant cell." Plant Molecular Biology Reporter 13.4 (1995): 355-362.

Wu, Y et al. "Event-specific qualitative and quantitative PCR detection methods for transgenic rapeseed hybrids MS1xRF1 and MS1xRF2" J. Agric. Food Chem. 55 (2007): 8380-8389.

Xu et al. "Event-specific detection of stacked genetically modified maize Bt11xGA21 by UP-M-PCR and real-time PCR" J. Agric. Food Chem. 57 (2009): 395-402.

Yu, Jianming, et al. "Genetic design and statistical power of nested association mapping in maize." Genetics 178.1 (2008): 539-551.

Žel, Jana, et al. How to reliably test for GMOs. Springer US, 2012. Henry, R. J. Molecular Markers in Plants. (Wiley, 2012).

Zhang, Changqing, et al. "Global characterization of cell-specific gene expression through fluorescence-activated sorting of nuclei." Plant physiology 147.1 (2008): 30-40.

---------- Page intentionally left blank ----------

European Commission
EUR 26705 EN – Joint Research Centre – Institute for Health and Consumer Protection

Title: **European Network of GMO Laboratories - Working Group for identification of stacked GM events (WG-IGSE)**
Working Group Report

Luxembourg: Publications Office of the European Union

2014 – 28 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series – ISSN 1831-9424 (online)

ISBN 978-92-79-38869-9 (PDF)

doi: 10.2788/87226

LB-NA-26705-EN-N

JRC Mission

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new methods, tools and standards, and sharing its know-how with the Member States, the scientific community and international partners.

Serving society
Stimulating innovation
Supporting legislation

doi: 10.2788/87226

ISBN 978-92-79-38869-9

Publications Office