European Commission

JRC TECHNICAL REPORTS

Using the EURL GMFF online bioinformatics resources

A How-To guide for practical JRC GMO-Matrix and JRC GMO-Amplicons case uses

Alexandre Angers
Mauro Petrillo
Peter Henriksson
Alex Patak

2016

Joint Research Centre

Using the EURL GMFF online bioinformatics resources

This publication is a Technical report by the Joint Research Centre, the European Commission's in-house science service. It aims to provide evidence-based scientific support to the European policy-making process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this publication.

Contact information
Molecular Biology and Genomics Unit
Joint Research Centre, Via Enrico Fermi 2749, 21027 Ispra (VA), Italy
E-mail: JRC-BGMO@ec.europa.eu
Tel.: +39 0332 789379

JRC Science Hub
https://ec.europa.eu/jrc

JRC101853

Table of contents

Abstract

The website of the European Reference Laboratory GM Food & Feed (EURL GMFF) (http://gmo-crl.jrc.ec.europa.eu) hosts a set of tools developed to assist laboratories involved in GMO monitoring.

These tools allow the mining of the information found in databases generated through bioinformatics analyses using the primer and probes information from the GMOMETHODS database, a database of reference methods for GMO analysis also hosted and maintained by the EURL GMFF.

The JRC GMO-Matrix, (http://gmo-crl.jrc.ec.europa.eu/jrcgmomatrix/), compiles predictions related to the detection of GM events by each of the detection methods, through *in silico* simulation of the Polymerase Chain Reaction results.

The JRC GMO-Amplicons, (http://gmo-crl.jrc.ec.europa.eu/jrcgmoamplicons/), allows access to a novel database of amplicon sequences generated by scanning publicly available sequence archives, in particular the poorly annotated patent sequences, with the detection methods used for GMO detection.

The aim of this document is to serve as a simple how-to guide for these applications, describing various practical case uses for which they were designed to provide support.

1. Introduction

The European Reference Laboratory GM Food & Feed (EURL GMFF) hosts a website where specific information can be found regarding, among others, the current status of the validation process for dossiers processed within the frame of the Regulation (EC) No 1829/2003 as well as various guidance documents. The EURL GMFF website also hosts the EU Database of Reference Methods for GMO Analysis (the GMOMETHODS database, see Bonfini et al, 2012), that includes complete information on the primer and probe sequences, as well as detailed performance indicators of methods for GMO analysis that have been either validated in a collaborative trial, according to the principles and requirements of ISO 5725 and/or IUPAC protocol, or verified by the EURL GMFF in the context of EU legislative acts.

The EURL GMFF also developed a set of tools aiming to provide information in support to the activities of testing laboratories. These include the JRC GMO-Matrix, hosted at http://gmo-crl.jrc.ec.europa.eu/jrcgmomatrix/, and the JRC GMO-Amplicons, hosted at http://gmo-crl.jrc.ec.europa.eu/jrcgmoamplicons/. Detailed description of their purpose and development can be found in their respective published articles (Angers-Loustau et al, 2014; Petrillo et al, 2015).

This technical report describes a set of procedures step by step in order to illustrate practical case uses and to explain the ways to obtain the appropriate information for each scenario.

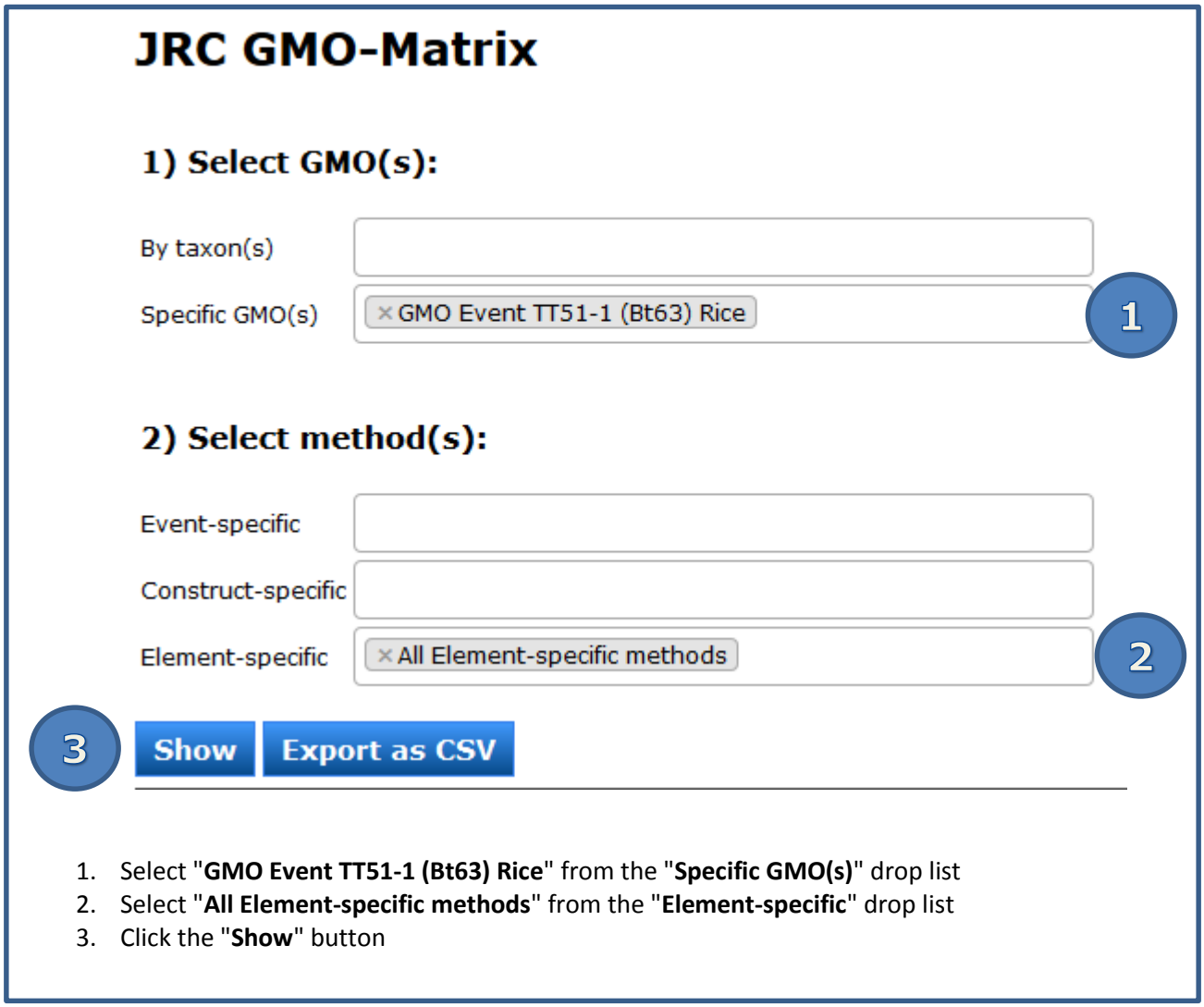
2. The JRC GMO-Matrix

2.1 Identifying the screening methods able to detect a specific GM event

It is sometimes necessary to detect a specific event in samples, when no event-specific method is available. In these cases, one possibility involves the use of a combination of screening methods that would detect different components of the event in question.

2.1.1 How to do it

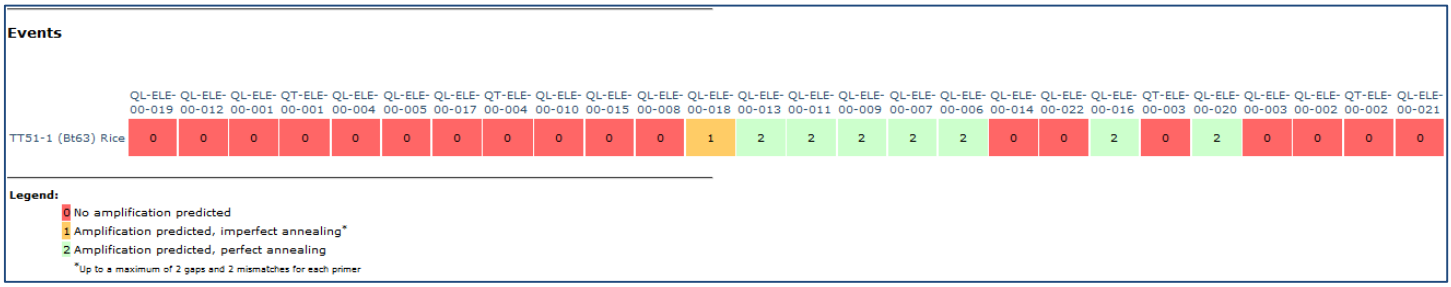
For this example, we will show how to obtain this information for the Bt63 rice event. The interface to use is the main matrix page:
http://gmo-crl.jrc.ec.europa.eu/jrcgmomatrix/matrices/full



1. Select "**GMO Event TT51-1 (Bt63) Rice**" from the "**Specific GMO(s)**" drop list
2. Select "**All Element-specific methods**" from the "**Element-specific**" drop list
3. Click the "**Show**" button

2.1.2 Interpreting the results

The result of this procedure is a matrix with a single row:



Events	QL-ELE-00-019	QL-ELE-00-012	QL-ELE-00-001	QT-ELE-00-001	QL-ELE-00-004	QL-ELE-00-005	QL-ELE-00-017	QT-ELE-00-004	QL-ELE-00-010	QL-ELE-00-015	QL-ELE-00-008	QL-ELE-00-018	QL-ELE-00-013	QL-ELE-00-011	QL-ELE-00-009	QL-ELE-00-007	QL-ELE-00-006	QL-ELE-00-014	QL-ELE-00-022	QL-ELE-00-016	QT-ELE-00-003	QL-ELE-00-020	QL-ELE-00-003	QL-ELE-00-002	QT-ELE-00-002	QL-ELE-00-021
TT51-1 (Bt63) Rice	0	0	0	0	0	0	0	0	0	0	0	1	2	2	2	2	2	0	0	2	0	2	0	0	0	0

Legend:

0 No amplification predicted
1 Amplification predicted, imperfect annealing*
2 Amplification predicted, perfect annealing
*Up to a maximum of 2 gaps and 2 mismatches for each primer

Each of the columns of the matrix represents the predicted amplification by the different element-specific methods in the GMOMETHODS database. As explained by the legend, a red (0) cell means that the method is expected to give a negative result when testing Bt63, while a green (2) cell means the methods will detect the event.

In this case, Bt63 will be detected by QL-ELE-00-013, QL-ELE-00-011, QL-ELE-00-009, QL-ELE-00-007, QL-ELE-00-006, QL-ELE-00-016 and QL-ELE-00-20.

QL-ELE-00-018, on the other hand, shows an orange (1) result, meaning that an amplicon was predicted, but some mismatches and/or gaps were found in the primers annealing regions. This might have an unpredictable impact on the amplification efficiency, and should be avoided if possible.

2.1.3 Additional considerations

To have more information on the methods present in the matrix, the header of the column is a link to the corresponding record in the GMOMETHODS database, where all the details (primer sequences, references, performance, etc.) can be found.

In addition to element-specific methods, the same procedure can be done with construct-specific methods, selecting "All Construct-specific methods" from the "Construct-specific" drop list in step 2.

In the case of stacked events, the individual events present in the stack should be selected in step 1, and the results should be combined by the user. Only methods giving a negative result for all events in the stacks will be negative for the whole stack.

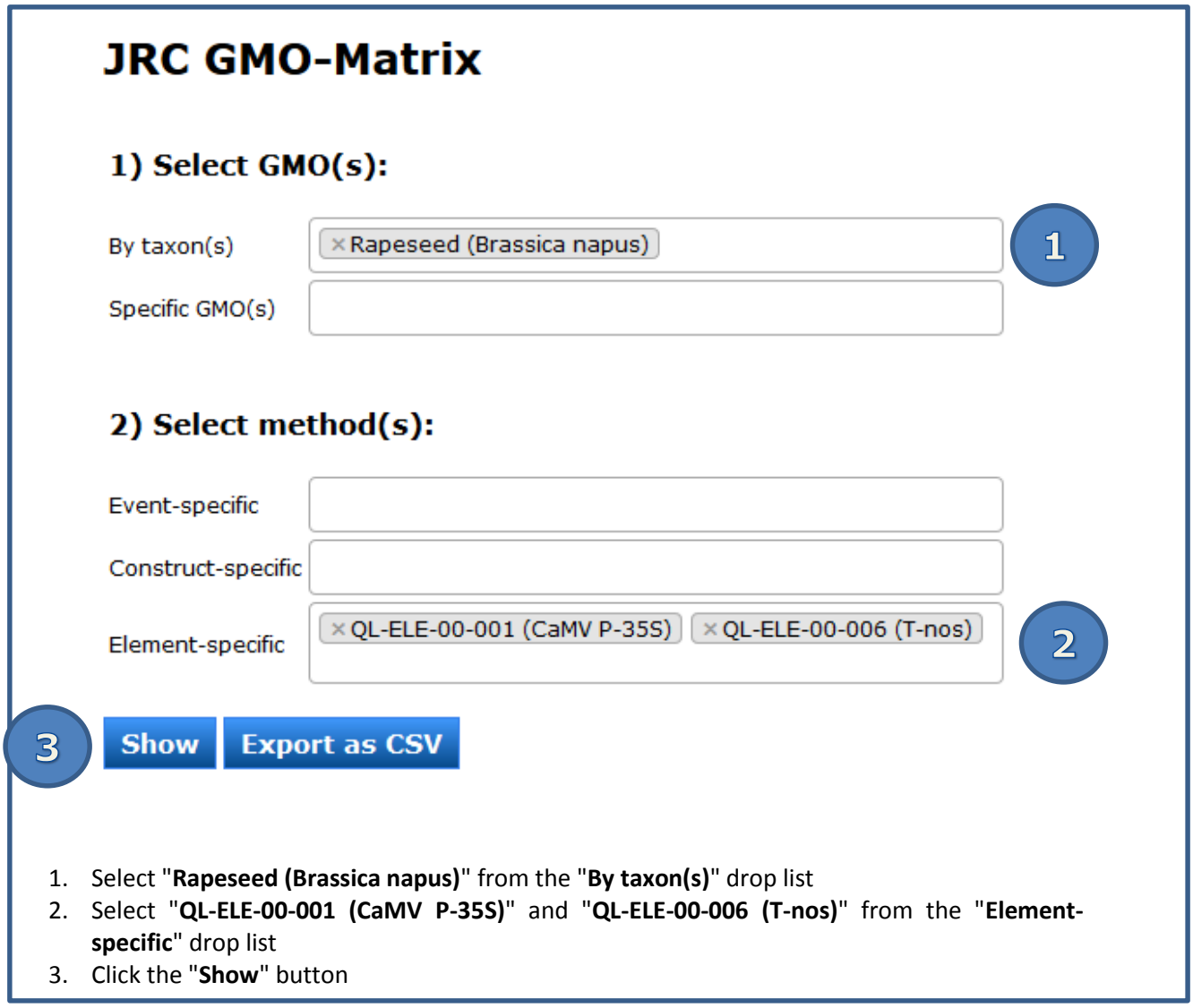
2.2 Evaluating the coverage of a testing strategy

When using a set of screening methods to routinely screen samples, it is important to know whether this combination of methods is able to detect all the known GM events in circulation for this taxon.

2.2.1 How to do it

For this example, we will test whether a combination of QL-ELE-00-001 (CaMV p35S) and QL-ELE-00-006 (T-nos) can detect all rapeseed events. The interface to use is the main matrix page:

http://gmo-crl.jrc.ec.europa.eu/jrcgmomatrix/matrices/full



1. Select "**Rapeseed (Brassica napus)**" from the "**By taxon(s)**" drop list
2. Select "**QL-ELE-00-001 (CaMV P-35S)**" and "**QL-ELE-00-006 (T-nos)**" from the "**Element-specific**" drop list
3. Click the "**Show**" button

2.2.2 Interpreting the results

The result of this procedure is a matrix with the different rapeseed events as rows, and the two selected methods as columns:

Events

	QL-ELE-00-001	QL-ELE-00-006
GT73 Rapeseed (MON-00073-7)	0	0
MS8 Rapeseed (ACS-BN005-8)	0	2
RF3 Rapeseed (ACS-BN003-6)	0	2
T45 Rapeseed (ACS-BN008-2)	2	0
Topas 19/2 Rapeseed (ACS-BN007-1)	2	0
Rf1 Rapeseed (ACS-BN001-4)	0	2
Ms1 Rapeseed (ACS-BN004-7)	0	2
Rf2 Rapeseed (ACS-BN002-5)	0	2
MON 88302 Rapeseed (MON-88302-9)	0	0
73496 Rapeseed (DP-073496-4)	0	0
Oxy-235 Rapeseed (ACS-BN011-5)	2	0

Legend:

0 No amplification predicted
1 Amplification predicted, imperfect annealing*
2 Amplification predicted, perfect annealing
*Up to a maximum of 2 gaps and 2 mismatches for each primer

Each of the cells of the matrix represents the predicted detection by the two methods of the different rapeseed events in the GMO-Matrix database. As explained by the legend, a red (0) cell means that the method is expected to give a negative result when testing that specific event, while a green (2) cell means the methods will detect the event.

The matrix shows that most of the rapeseed events will be detected either by the QL-ELE-00-001 (CaMV p35S) or the QL-ELE-00-006 (T-nos) method, with none of them being detected by both. More importantly, three of these events, GT73 Rapeseed (MON-00073-7), MON 88302 Rapeseed (MON-88302-9) and 73496 Rapeseed (DP-073496-4), are detected by neither the methods. If detecting these events is considered to be important in the monitoring framework, additional methods will be needed.

2.2.3 Additional considerations

If needed, more than one taxon can be selected in step 1

Information on the different rapeseed events can be found at the linked record in the BCH LMO Registry[1] by clicking on the name of the event in the first column.

1 https://bch.cbd.int/database/lmo-registry/

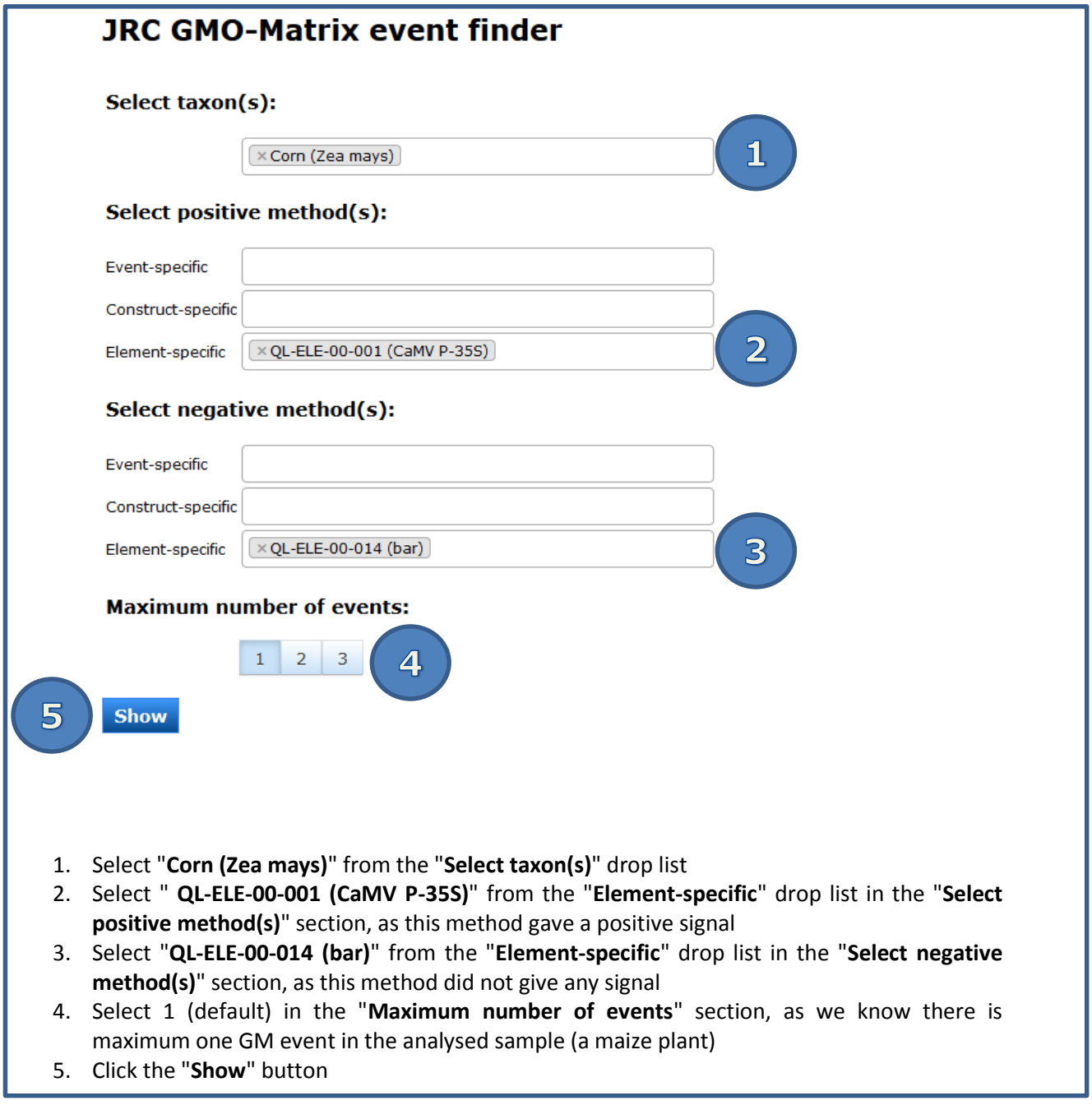
2.3 Analysing the results of a set of screening tests

After performing a set of screening experiments using methods from the GMOMETHODS database, it is possible to use the JRC GMO-Matrix to identify the different GM events that are compatible with the pattern of results obtained.

2.3.1 How to do it

For this example, we will analyse the following scenario: a maize plant was analysed with QL-ELE-00-001 (CaMV p35S) and QL-ELE-00-014 (bar). The CaMV p35S method was positive, while the bar method gave a negative result.

The interface to use is the Event finder (http://gmo-crl.jrc.ec.europa.eu/jrcgmomatrix/matrices/event_finder)



1. Select "**Corn (Zea mays)**" from the "**Select taxon(s)**" drop list
2. Select " **QL-ELE-00-001 (CaMV P-35S)**" from the "**Element-specific**" drop list in the "**Select positive method(s)**" section, as this method gave a positive signal
3. Select "**QL-ELE-00-014 (bar)**" from the "**Element-specific**" drop list in the "**Select negative method(s)**" section, as this method did not give any signal
4. Select 1 (default) in the "**Maximum number of events**" section, as we know there is maximum one GM event in the analysed sample (a maize plant)
5. Click the "**Show**" button

2.3.2 Interpreting the results

The main result of this procedure is a matrix with different corn events as rows, and the two selected methods as columns:

Single events:

	QL-ELE-00-001	QL-ELE-00-014
TC1507 Maize (DAS-01507-1)	2	0
59122 Maize (DAS-59122-7)	2	0
MON810 Maize (MON-00810-6)	2	0
MON863 Maize (MON-00863-5)	1	0
MON88017 Maize (MON-88017-3)	2	0
NK603 Maize (MON-00603-6)	2	0
T25 Maize (ACS-ZM003-2)	2	0
MON89034 Maize (MON-89034-3)	2	0
Bt11 Maize (SYN-BT011-1)	2	0
32 Maize (DAS-59132-8)	2	0
MON87460 Maize (MON-87460-4)	1	0
MON 87427 Maize (MON-87427-7)	2	0
4114 Maize (DP-004114-3)	2	0
MON 87411 Maize (MON-87411-9)	2	0
32316 Maize (DP-032316-8)	2	0

Legend:

0 No amplification predicted
1 Amplification predicted, imperfect annealing*
2 Amplification predicted, perfect annealing

*Up to a maximum of 2 gaps and 2 mismatches for each primer

Each of the cells of the matrix represents the predicted detection by the two methods of the shown maize events. As explained by the legend, a red (0) cell means that the method is expected to give a negative result when testing that specific event, while a green (2) cell means the methods will detect the event. An orange (1) result is considered for these analyses as a positive signal, meaning that an amplicon was predicted, but some mismatches and/or gaps were found in the primer annealing regions.

This matrix only shows the events in the GMO-Matrix database that match the patterns selected in the form: 1) events of maize origin, 2) that are expected to be positive for CaMV P-35S and 3) negative for bar. In this example, fifteen events fit that pattern. Most probably, the sample contains one of these GM events, unless another GMO(s), not present in the JRC GMO-Matrix database, is present.

2.3.3 Additional considerations

Samples of unknown composition

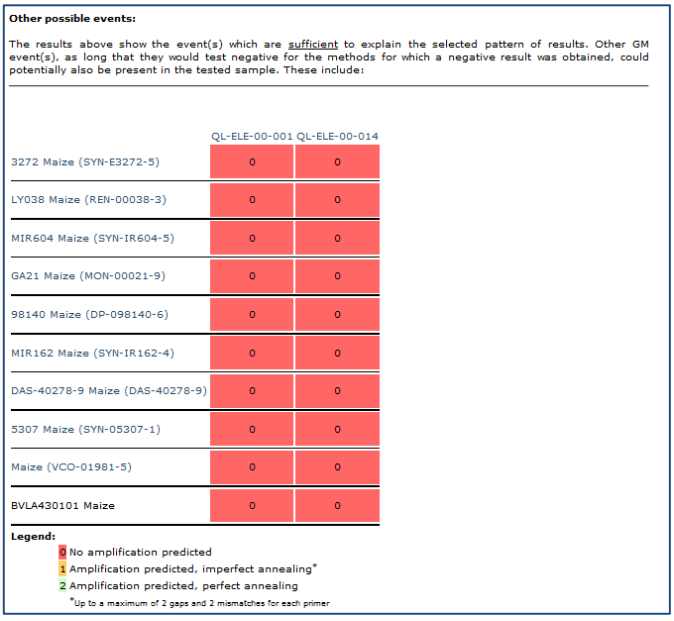
In the previous example, the taxon of the tested sample was known (maize). However, if this is not the case, the species present in the sample need to be determined, as a first step, using a set of taxon-specific methods such as those present in the GMOMETHODS database. Then, all the taxons detected can be selected together in step 1

GMO masking

The results presented by the Event Finder interface show the events which are sufficient to explain the selected pattern of positive/negative results.

However, more than one GM event can be present, especially in samples composed of a mixture of crops. In particular, in addition to one or more of the events suggested in the matrix describe above, other GM event(s) could potentially also be present in the tested sample if they would test negative for the methods selected.

The interface presents a list of these events in a separate matrix, below the first one, showing the events present in the JRC GMO-Matrix that fit the criteria:



Other possible events:

The results above show the event(s) which are sufficient to explain the selected pattern of results. Other GM event(s), as long that they would test negative for the methods for which a negative result was obtained, could potentially also be present in the tested sample. These include:

	QL-ELE-00-001	QL-ELE-00-014
3272 Maize (SYN-E3272-5)	0	0
LY038 Maize (REN-00038-3)	0	0
MIR604 Maize (SYN-IR604-5)	0	0
GA21 Maize (MON-00021-9)	0	0
98140 Maize (DP-098140-6)	0	0
MIR162 Maize (SYN-IR162-4)	0	0
DAS-40278-9 Maize (DAS-40278-9)	0	0
5307 Maize (SYN-05307-1)	0	0
Maize (VCO-01981-5)	0	0
BVLA430101 Maize	0	0

Legend:

0 No amplification predicted
1 Amplification predicted, imperfect annealing*
2 Amplification predicted, perfect annealing

*Up to a maximum of 2 gaps and 2 mismatches for each primer

Finding additional methods to refine the candidate list

In the example shown, two screening methods narrowed the list of candidate GM events to 15 events down from the 26 maize events present in the JRC GMO-Matrix at the time this How-To guide was made.

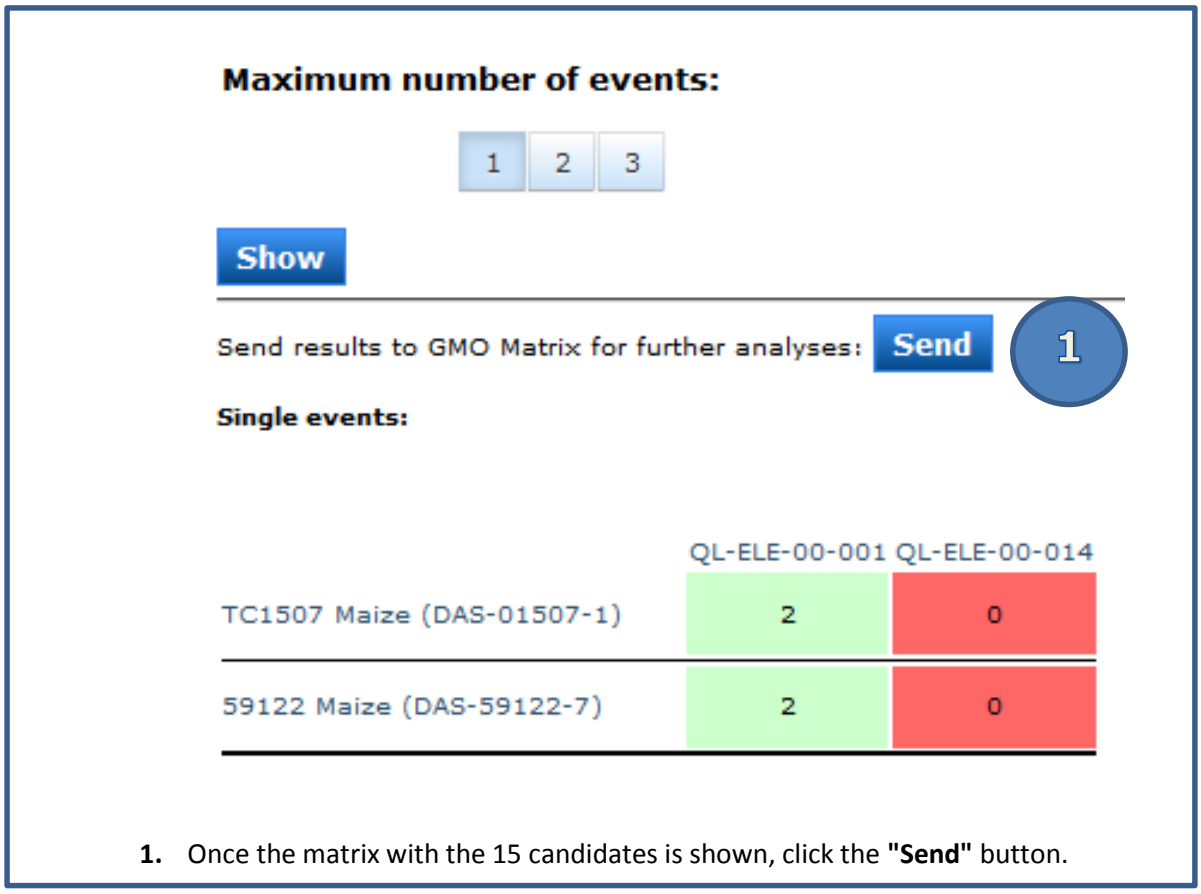
The next step of the analyses is then an attempt to perform further testing, with screening methods, to reduce the list of candidate GM events. One way to do this, using the JRC GMO-Matrix, is shown in the next section.

2.4 Choosing methods for an efficient screening strategy

Section 2.3 showed the results of an analysis using two screening methods that produced a list of 15 candidate GM events. It is possible to use the JRC GMO-Matrix to identify methods that would allow to efficiently narrow down the list of potential GM event(s) present in the sample.

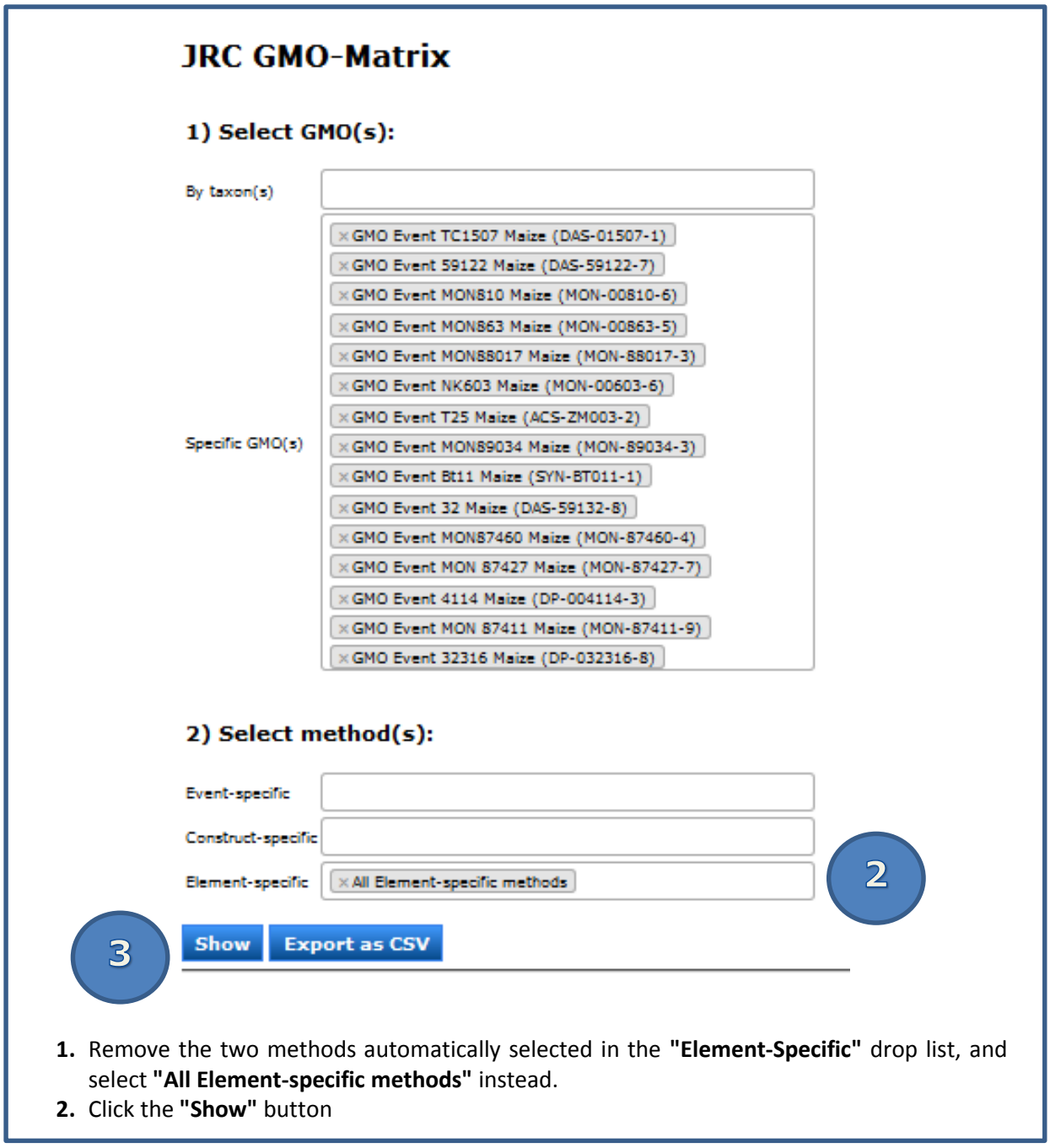
2.4.1 How to do it

For this example, we will continue the analysis from section 2.3. If not already done, perform the steps described in section 2.3.1.



1. Once the matrix with the 15 candidates is shown, click the **"Send"** button.

This will open the JRC GMO-Matrix, automatically selecting the 15 events identified by the Event Finder interface.



1. Remove the two methods automatically selected in the **"Element-Specific"** drop list, and select **"All Element-specific methods"** instead.
2. Click the **"Show"** button

2.4.2 Interpreting the results

Events

	QL-ELE-00-019	QL-ELE-00-012	QL-ELE-00-001	QT-ELE-00-001	QL-ELE-00-004	QL-ELE-00-005	QL-ELE-00-017	QT-ELE-00-004	QL-ELE-00-010	QL-ELE-00-015	QL-ELE-00-008	QL-ELE-00-018	QL-ELE-00-013	QL-ELE-00-011	QL-ELE-00-009	QL-ELE-00-007	QL-ELE-00-006	QL-ELE-00-014	QL-ELE-00-022	QL-ELE-00-016	QT-ELE-00-003	QL-ELE-00-020	QL-ELE-00-003	QL-ELE-00-002	QT-ELE-00-002	QL-ELE-00-021
TC1507 Maize (DAS-01507-1)	0	2	2	2	2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2
59122 Maize (DAS-59122-7)	0	2	2	2	2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2
MON810 Maize (MON-00810-6)	0	2	2	2	2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
MON863 Maize (MON-00863-5)	0	2	1	1	2	1	2	2	0	0	0	1	2	2	2	2	2	0	0	0	0	0	2	2	0	0
MON88017 Maize (MON-88017-3)	1	2	2	2	2	2	2	2	0	0	0	1	2	2	2	2	2	0	0	0	0	0	0	0	0	0
NK603 Maize (MON-00603-6)	1	2	2	2	2	2	2	2	0	0	0	1	2	2	2	2	2	0	0	0	0	0	0	0	0	0
T25 Maize (ACS-ZM003-2)	0	2	2	2	2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2
MON89034 Maize (MON-89034-3)	0	2	2	2	2	2	2	2	2	2	0	1	2	2	2	2	2	0	0	0	0	1	0	0	0	0
Bt11 Maize (SYN-BT011-1)	0	2	2	2	2	2	2	2	0	0	0	1	2	2	2	2	2	0	0	2	0	2	0	0	2	2
32 Maize (DAS-59132-8)	0	2	2	2	2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2
MON87460 Maize (MON-87460-4)	0	2	1	1	2	1	2	2	0	0	0	1	2	2	2	2	2	0	0	0	0	0	2	2	0	0
MON 87427 Maize (MON-87427-7)	0	2	2	2	2	2	2	2	0	0	0	1	2	2	2	2	2	0	0	0	0	0	0	0	0	0
4114 Maize (DP-004114-3)	0	2	2	2	2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2
MON 87411 Maize (MON-87411-9)	1	2	2	2	2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32316 Maize (DP-032316-8)	0	2	2	2	2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2

Legend:

0 No amplification predicted
1 Amplification predicted, imperfect annealing*
2 Amplification predicted, perfect annealing

*Up to a maximum of 2 gaps and 2 mismatches for each primer

Each of the cells of the matrix represents the predicted detection by the element-specific methods of the maize events identified by the Event finder step. As explained by the legend, a red (0) cell means that the method is expected to give a negative result when testing that specific event, while a green (2) cell means the methods will detect the event. An orange (1) result means that an amplicon was predicted, but some mismatches and/or gaps were found in the primer annealing regions.

This matrix shows that two groups of element-specific methods in the GMOMETHODS database could be good candidates for the subsequent screening step, as they split the candidate events into two groups of almost equal numbers: QL-ELE-00-018, -013, -011, -009, -007 and -006 that target the nopaline synthase terminator (T-nos), as well as the QT-ELE-002 and QL-ELE-00-021 methods, targeting the phosphinothricin N-acetyltransferase (pat) gene.

2.4.3 Additional considerations

This type of analyses can also be performed prior to any testing, selecting whole taxons in the JRC GMO-Matrix form instead of specific events.

2.5 Analysing the results of a Pre-Spotted Plate (PSP) experiment

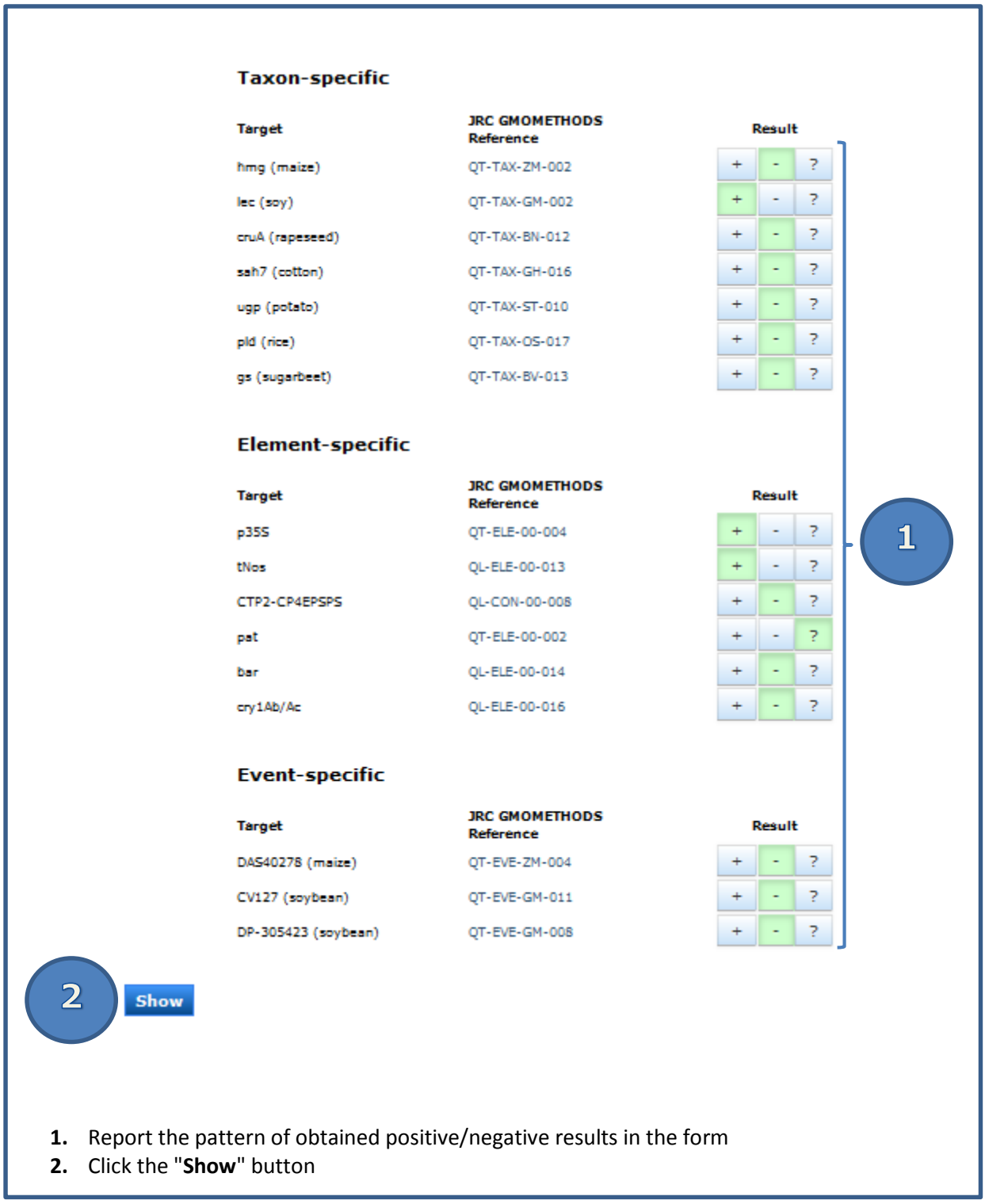
Efficient screening strategies can be developed using various panels of element- and construct-specific methods. The Molecular Biology and Genomics (MBG) Unit of the JRC Institute for Health and Consumer Protection has developed a ready-to-use, multi-target screening tool known as the pre-spotted plate (PSP) that allows testing for the presence of multiple GMO-targets in a single PCR experiment. A specific interface, with a defined set of methods, was developed to analyse the results obtained using these plates.
The interface is available at
http://gmo-crl.jrc.ec.europa.eu/jrcgmomatrix/matrices/prespotted_plates

2.5.1 How to do it

For this example, we will assume the PSP experiment produced the following results:

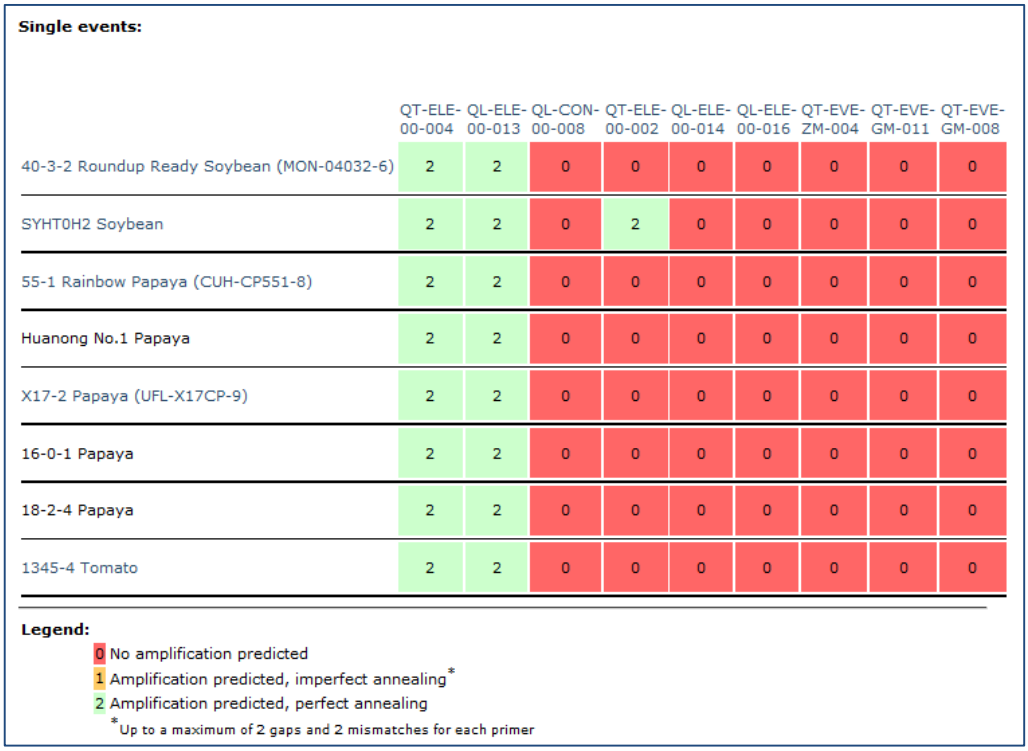
Target	JRC GMOMETHODS Reference	Result
hmg (maize)	QT-TAX-ZM-002	-
lec (soy)	QT-TAX-GM-002	+
cruA (rapeseed)	QT-TAX-BN-012	-
sah7 (cotton)	QT-TAX-GH-016	-
ugp (potato)	QT-TAX-ST-010	-
pld (rice)	QT-TAX-OS-017	-
gs (sugarbeet)	QT-TAX-BV-013	-
Target	**JRC GMOMETHODS Reference**	**Result**
p35S	QT-ELE-00-004	+
tNos	QL-ELE-00-013	+
CTP2-CP4EPSPS	QL-CON-00-008	-
pat	QT-ELE-00-002	Ambiguous*
bar	QL-ELE-00-014	-
cry1Ab/Ac	QL-ELE-00-016	-
Target	**JRC GMOMETHODS Reference**	**Result**
DAS40278 (maize)	QT-EVE-ZM-004	-
CV127 (soybean)	QT-EVE-GM-011	-
DP-305423 (soybean)	QT-EVE-GM-008	-

* +/- result



1. Report the pattern of obtained positive/negative results in the form
2. Click the "**Show**" button

2.5.2 Interpreting the results



Single events:

	QT-ELE-00-004	QL-ELE-00-013	QL-CON-00-008	QT-ELE-00-002	QL-ELE-00-014	QL-ELE-00-016	QT-EVE-ZM-004	QT-EVE-GM-011	QT-EVE-GM-008
40-3-2 Roundup Ready Soybean (MON-04032-6)	2	2	0	0	0	0	0	0	0
SYHT0H2 Soybean	2	2	0	2	0	0	0	0	0
55-1 Rainbow Papaya (CUH-CP551-8)	2	2	0	0	0	0	0	0	0
Huanong No.1 Papaya	2	2	0	0	0	0	0	0	0
X17-2 Papaya (UFL-X17CP-9)	2	2	0	0	0	0	0	0	0
16-0-1 Papaya	2	2	0	0	0	0	0	0	0
18-2-4 Papaya	2	2	0	0	0	0	0	0	0
1345-4 Tomato	2	2	0	0	0	0	0	0	0

Legend:

0 No amplification predicted
1 Amplification predicted, imperfect annealing*
2 Amplification predicted, perfect annealing

*Up to a maximum of 2 gaps and 2 mismatches for each primer

Each of the cells of the matrix represents the predicted detection by the PSP methods of the shown events. As explained by the legend, a red (0) cell means that the method is expected to give a negative result when testing that specific event, while a green (2) cell means the methods will detect the event. An orange (1) result is considered for these analyses as a positive signal, meaning that an amplicon was predicted, but some mismatches and/or gaps were found in the primer annealing regions.

When one of the methods in the PSP shows an ambiguous result (as defined in the protocol describing the use of these plates), the tool will present events that either contain or do not contain the element, as is the case in the shown example with the QT-ELE-00-002 method.

This matrix only shows the events in the GMO-Matrix database that match the patterns selected in the form. In this example, 8 events fit that pattern.

2.5.3 Additional considerations

Even though the only taxon-specific method that was positive in the PSP was soybean, in the results shown in this example only two of the eight GM events identified are of soybean origin, while the others are of different species (papaya and tomato). This is because the PSP currently does not contain taxon-specific methods that can detect the presence/absence of these species in the sample, so the tool can't eliminate the possibility that they are present, and shows them in the result.

GMO masking
Since this interface is a specialised version of the Event Finder interface, the same caveat holds for the concept of GMO masking when analysing the results of a screening experiment. Refer to section 2.3.3 for more details.

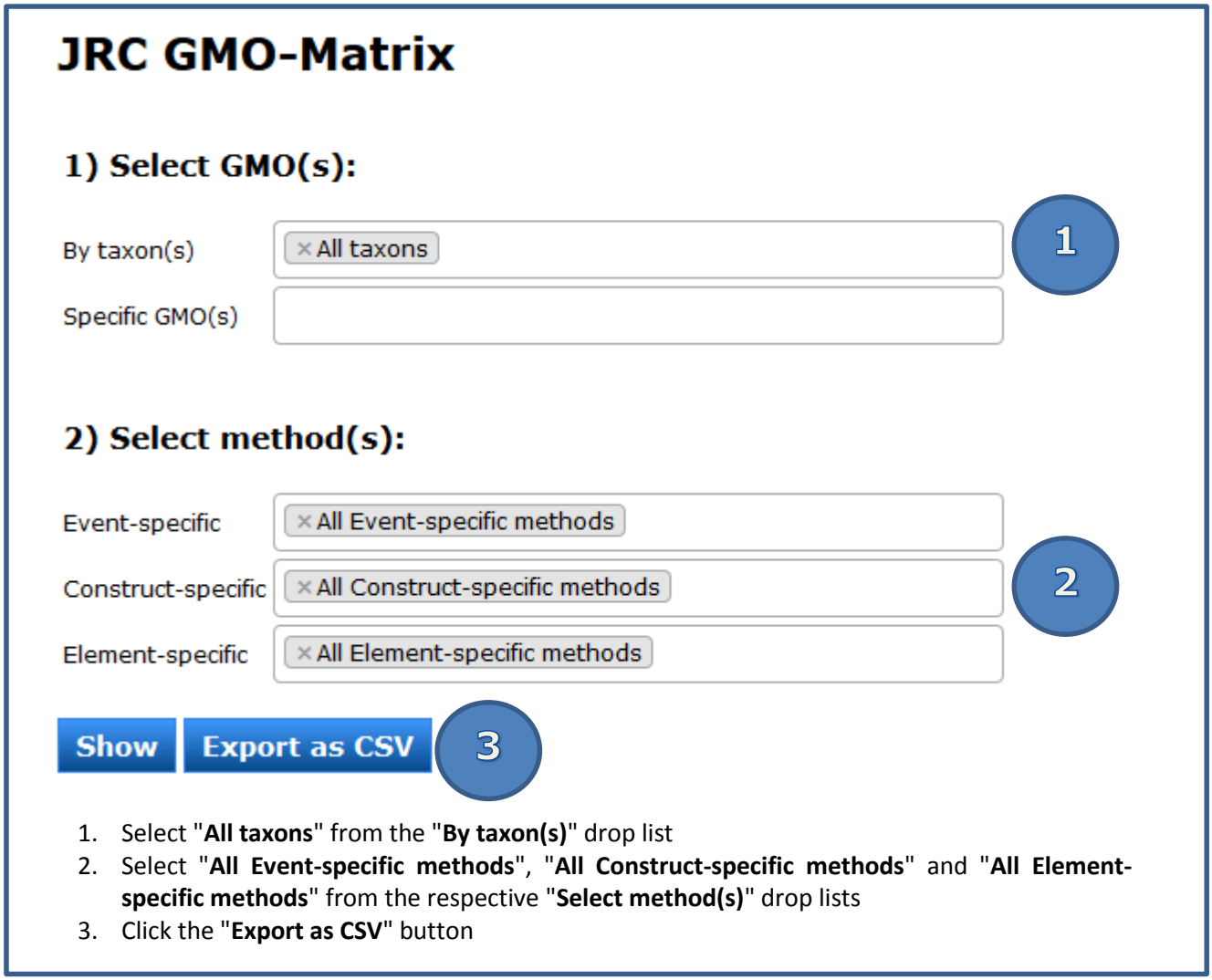
2.6 Downloading the whole JRC GMO-Matrix database to Excel

The web portal of the JRC GMO-Matrix provides a set of tools that are used to mine the data contained in its database, as described in the previous sections.

However, it is also possible to export a local copy of all the information found in the database and manipulate it in a program such as Excel for purposes or needs not covered in the supplied interface.

2.6.1 How to do it

The interface to use is the main matrix page (http://gmo-crl.jrc.ec.europa.eu/jrcgmomatrix/matrices/full)



1. Select "**All taxons**" from the "**By taxon(s)**" drop list
2. Select "**All Event-specific methods**", "**All Construct-specific methods**" and "**All Element-specific methods**" from the respective "**Select method(s)**" drop lists
3. Click the "**Export as CSV**" button

2.6.2 Interpreting the results

Once the "Export as CSV" button is clicked, the JRC GMO-Matrix will produce a file that will be downloaded by the internet browser and saved locally.

This file is a text file containing the information of the JRC GMO-Matrix database, in the "comma-separated values" (CSV) format.

This file can be imported in spreadsheet programs (such as Excel), using the comma as a column separator character. The result is a large matrix contain all the GM events as rows, and all the GMOMETHODs methods as columns. The matrix can then be manipulated as a normal spreadsheet.

2.6.3 Additional considerations

It is not necessary to export the content of the whole JRC GMO-Matrix database. It is possible to select subsets of events or taxons (step 1), or subsets of methods (step 2), and the resulting file will be constructed accordingly.

3. The JRC GMO-Amplicons

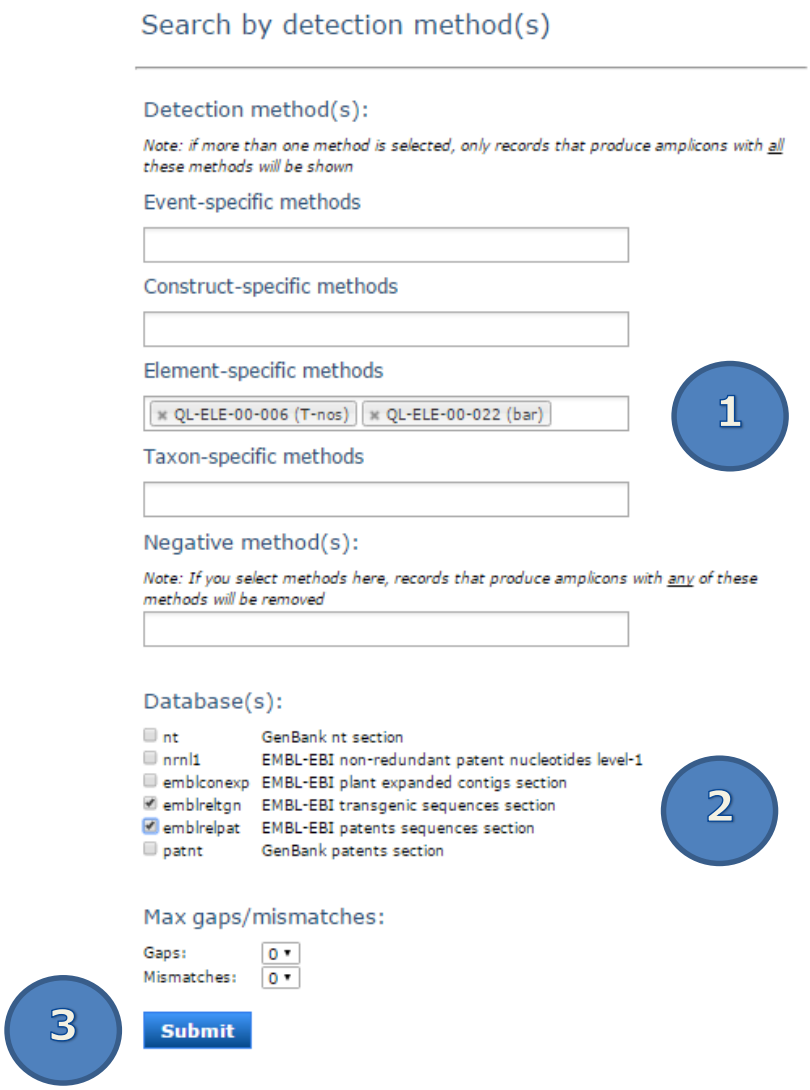
3.1 Investigating an unexpected screening pattern

The GM events present in the JRC GMO-Matrix database are those in which complete transgenic insert sequences are present in the internal Central Core Sequence Information System of the Joint Research Centre (Patak, 2011). When the pattern of a set of screening tests is not compatible with any of these events, it is possible to scan the database of the JRC GMO-Amplicons to verify whether sequences from other public databases match the screening result.
This tool aims at identifying "unknown" GMOs for which some sequence information is publicly available.

3.1.1 How to do it

For this example, we will assume that a maize sample was found positive for the following two methods: QL-ELE-00-006 (T-nos) and QL-ELE-00-022 (bar).
Using the Event Finder of the JRC GMO-Matrix interface, as explained in section 2.3, potential cotton, rapeseed and rice events (but none in maize) are identified.

The interface to use is the **Amplicon Finder** form
http://gmo-crl.jrc.ec.europa.eu/jrcgmoamplicons/db_scans/form)

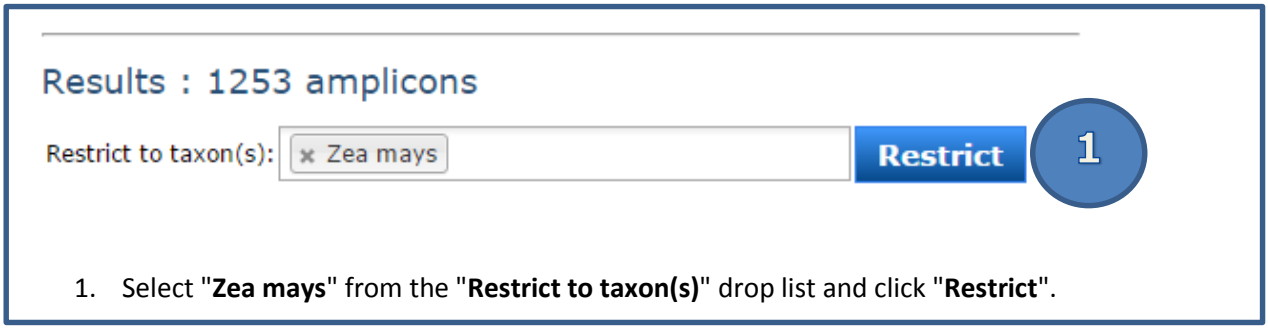


1. Select "**QL-ELE-00-006 (T-nos)**" and "**QL-ELE-00-022 (bar)**" from the "**Element-specific methods**" drop list
2. Select "**emblretgn**" and "**emblrelpat**" from the "**Databases(s)**" list
3. Click the "**Submit**" button

3.1.2 Interpreting the results

The JRC GMO-Amplicons populates a table with a list of all the records in the selected transgenic and patents sequence databases that would produce a pattern of amplifications matching the methods entered in the form. This list, at the time of this manual, contains 1253 amplicons.

It is possible, at this time, to further refine the list by taxon, as the sample tested was maize.



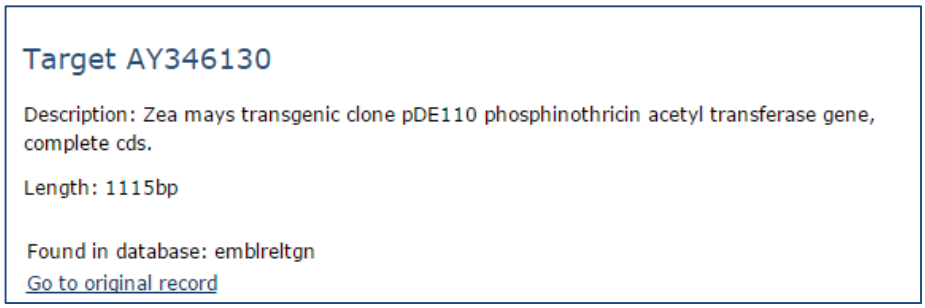
1. Select "**Zea mays**" from the "**Restrict to taxon(s)**" drop list and click "**Restrict**".

The resulting table now contains a single record. Since the JRC GMO-Amplicons is a database of amplicons, there are two entries in the final table, corresponding to the two methods selected (QL-ELE-00-006 and QL-ELE-00-022):

Record	DB	Record description	Record length	Taxon	Method	Amplicon size	Match details
AY346130	emblreltgn	Zea mays transgenic clone pDE110 ph...	1115	Zea mays	QL-ELE-00-006	180bp	View
AY346130	emblreltgn	Zea mays transgenic clone pDE110 ph...	1115	Zea mays	QL-ELE-00-022	69bp	View

Clicking on **"View"** in the *Match details* column allows seeing the predicted sequences of the amplicons produced by the two methods on this record.

Clicking on the record ID (*e.g.* AY346130) in the *Record* column shows more information on the target sequence itself. The details of the record page will be described in more details in section 3.2 below, but for now the interesting information is at the top of the page:



The whole description of the record shows that the pattern of screening results is consistent with the *Zea mays* transgenic clone pDE110. Following the "**Go to original record**" link opens the record page in the ENA sequence database, where the following reference is found: "Development of real-time PCR method for quantification of maize transgenic line StarLink".

Together, this information suggests that "Starlink" (which is not in the JRC GMO-Matrix database) is a good candidate for the detected maize GM event. This is consistent with additional information available for this event, for example in the Biosafety Scanner

website, that shows that the construct for this event contains both a bar gene and the T-nos promoter (http://en.biosafetyscanner.org/schedaevento.php?evento=109).

3.1.3 Additional considerations

The form to select the methods can also include methods that were negative in the screening experiment, in the "**Negative method(s)**" droplist. In this single list are included Event-, Construct-, Element- and Taxon- specific methods.

By default, the JRC GMO-Amplicons will only select records with perfect annealing of probes and primers. This can be relaxed to allow some gaps and mismatches (maximum 2) in the "**Max gaps/mismatches**" section.

The "**Restrict to taxon(s)**" list is populated for each request, containing only the taxons reported for the identified records. If the species of interest is not present, it means that no record in the set matching the detection methods pattern was reported as this species. It should be kept in mind that for many of the records, in particular in the patents databases, the taxon is not specified, and taxons such as "synthetic constructs", "unidentified" and "artificial sequences" should also be considered.

3.2 Identifying methods able to detect a specific record

When a specific record of interest is found in the public sequence databases, it is possible to use the JRC GMO-Amplicons application to determine which detection methods would be able to detect it.

3.2.1 How to do it

For this example, we will analyse the pVE143 plasmid, described in the patent WO9213957, and found in the NCBI GenBank database with the accession number A23633.1 (http://www.ncbi.nlm.nih.gov/nuccore/836589).

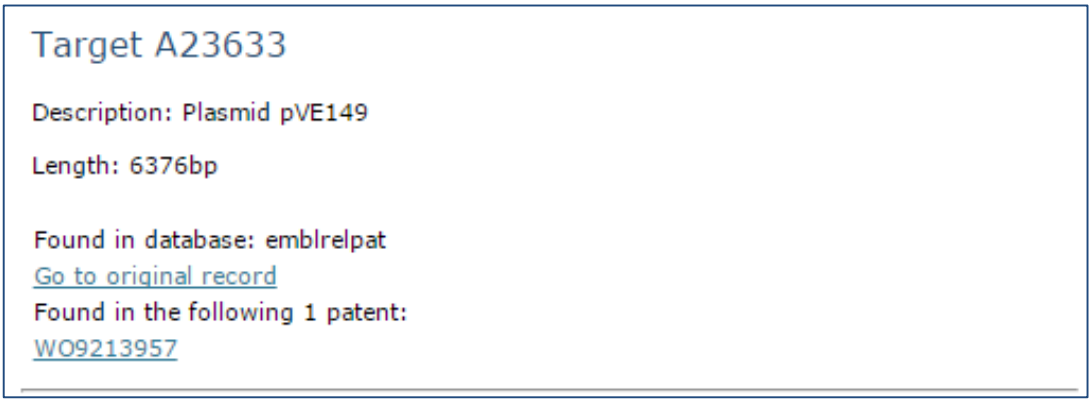
The interface to use is the "**Search by target ID**" form available at http://gmo-crl.jrc.ec.europa.eu/jrcgmoamplicons/db_scans/record

1. Enter the record ID, "**A23633**" in the "**Target ID**" field
2. Click the "**Show**" button

3.2.2 Interpreting the results

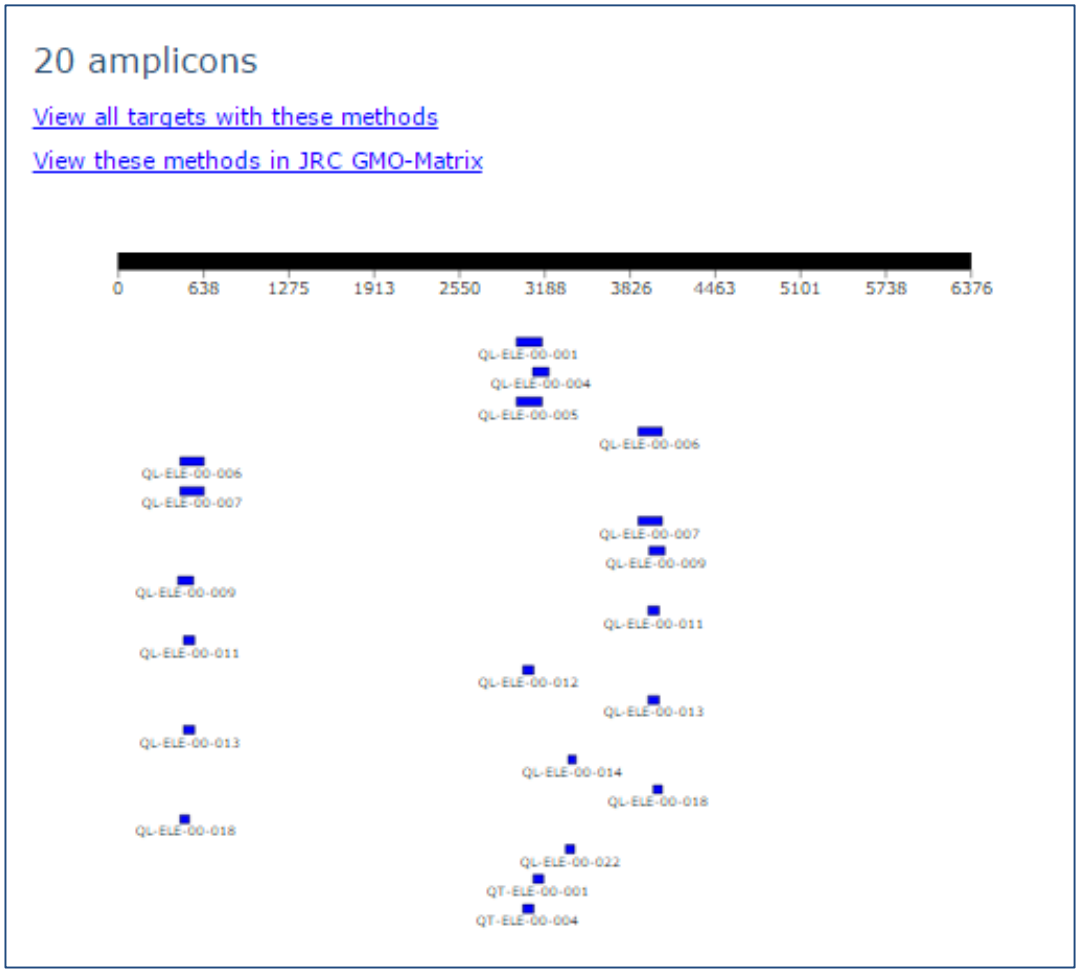
If a record is found in the JRC GMO-Amplicons database, the page will refresh with information about the record.

The first section contains the description of the record, its length, a link to the original record in the public databases, as well as a link to the associated patent in The Lens website[2] , if available:



[2] www.lens.org

The second section contains a graphical representation of the different amplicons predicted (with the methods from the JRC GMOMETHODS database) for this record and their position on the sequence of the record:



Each of the blue rectangles links to the detailed information about the respective amplicon sequence. This view shows that many methods are able to detect this record, some producing 2 independent amplicons (e.g. QL-ELE-00-011).

The last section is a table showing the same information in a tabular format.

3.2.3 Additional considerations

The database for JRC GMO-Amplicons is updated regularly, every six months on average. It is then possible that there is a delay between the time a new record is released in the public databases and its appearance in the JRC GMO-Amplicons. The exact date of the last update is shown on the last line of every page.

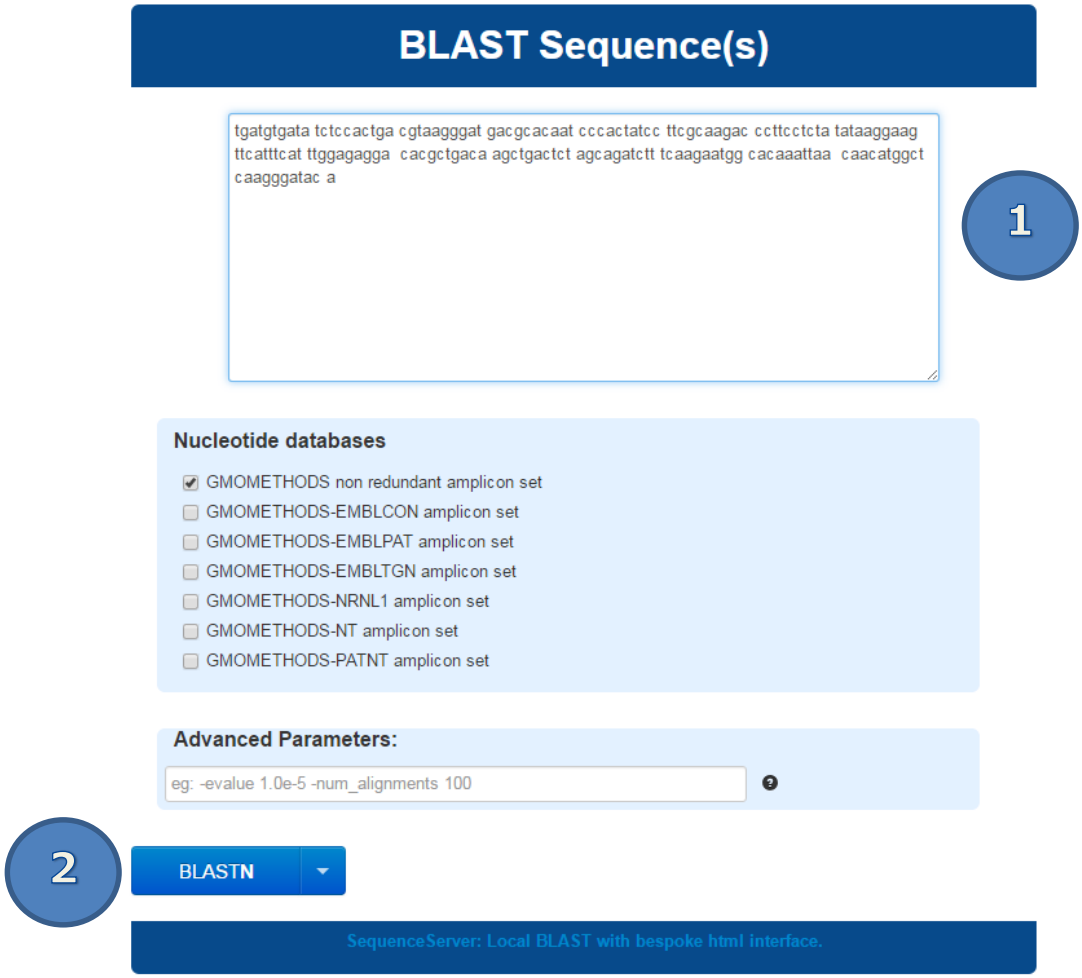
If a record older than that date does not give any result ("**No information found for record ...**"), it means that none of the methods in the JRC GMOMETHODS database were found to produce an amplicon on its sequence.

3.3 Searching the JRC GMO-Amplicons database with BLAST

The database of the JRC GMO-Amplicons application is, at its core, a DNA sequence database. It is possible to search through this database using BLAST.

3.3.1 How to do it

The interface to use is the "**BLAST amplicon**" form available at http://gmo-crl.jrc.ec.europa.eu/jrcgmoamplicons/db_scans/blast .

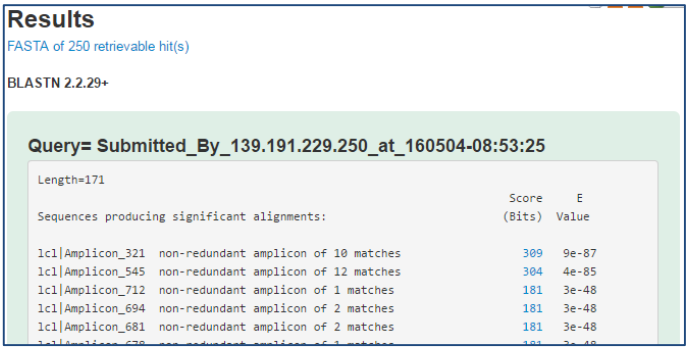


1. Enter the DNA sequence to be analysed in the text field
2. Click the "**BLASTN**" button

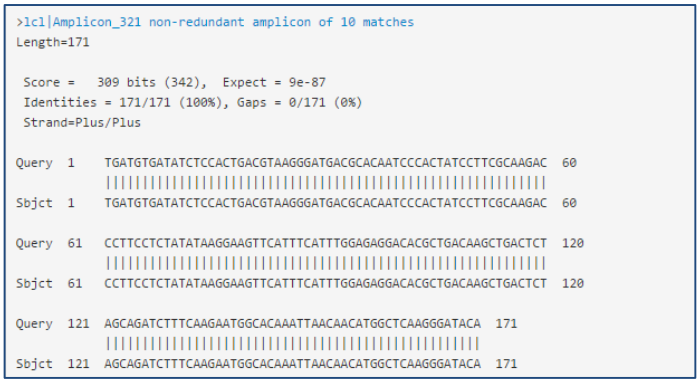
3.3.2 Interpreting the results

The JRC GMO-Amplicon will perform a BLAST search of the input sequence against the selected set of sequences, in that case the non-redundant amplicon set, selected by default. In this set, all identical sequences are combined into a single representative record.

The output is a standard BLAST report:

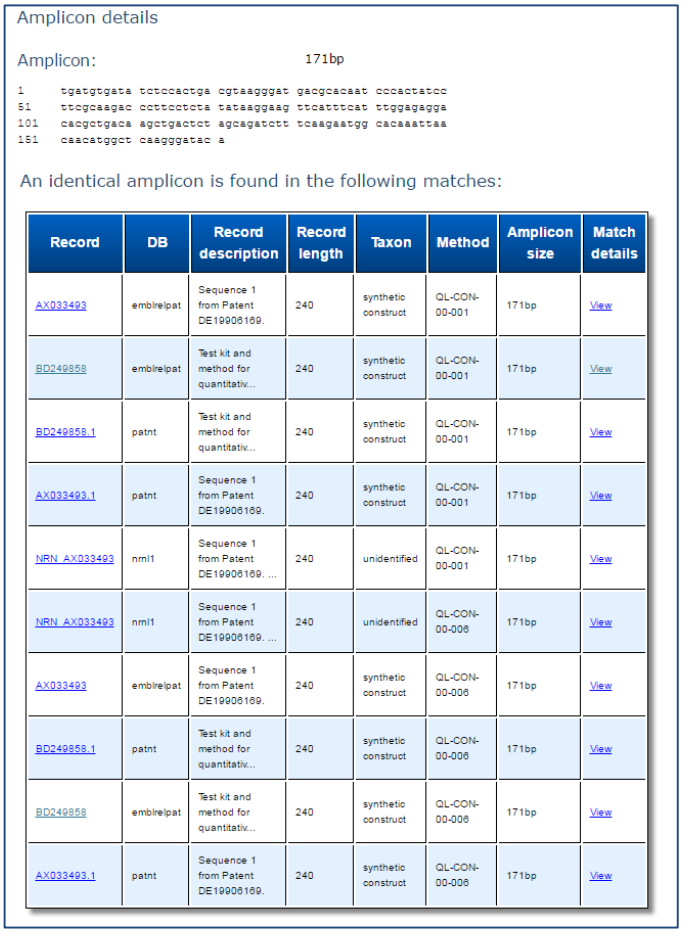


Clicking on the score of the best hit (309) scrolls the page to the details of the match:



The name of the hit states that this unique amplicon was found, identical, in 10 records of the JRC GMO-Amplicons database, and that it is perfect (171/171, with no gaps).

Following the link with the name of the amplicon brings a page showing all the records that contain this sequence:



Amplicon details

Amplicon: 171bp

```
1     tgatgtgata tctccactga cgtaagggat gacgcacaat cccactatcc
51    ttcgcaagac ccttcctcta tataaggaag ttcatttcat ttggagagga
101   cacgctgaca agctgactct agcagatctt tcaagaatgg cacaaattaa
151   caacatggct caagggatac a
```

An identical amplicon is found in the following matches:

Record	DB	Record description	Record length	Taxon	Method	Amplicon size	Match details
AX033493	emblrelpat	Sequence 1 from Patent DE19906169.	240	synthetic construct	QL-CON-00-001	171bp	View
BD249858	emblrelpat	Test kit and method for quantitativ...	240	synthetic construct	QL-CON-00-001	171bp	View
BD249858.1	patnt	Test kit and method for quantitativ...	240	synthetic construct	QL-CON-00-001	171bp	View
AX033493.1	patnt	Sequence 1 from Patent DE19906169.	240	synthetic construct	QL-CON-00-001	171bp	View
NRN_AX033493	nrnl1	Sequence 1 from Patent DE19906169. ...	240	unidentified	QL-CON-00-001	171bp	View
NRN_AX033493	nrnl1	Sequence 1 from Patent DE19906169. ...	240	unidentified	QL-CON-00-006	171bp	View
AX033493	emblrelpat	Sequence 1 from Patent DE19906169.	240	synthetic construct	QL-CON-00-006	171bp	View
BD249858.1	patnt	Test kit and method for quantitativ...	240	synthetic construct	QL-CON-00-006	171bp	View
BD249858	emblrelpat	Test kit and method for quantitativ...	240	synthetic construct	QL-CON-00-006	171bp	View
AX033493.1	patnt	Sequence 1 from Patent DE19906169.	240	synthetic construct	QL-CON-00-006	171bp	View

Each of the rows in this table links to the record and match details pages described previously, that give all the available information on the nature of the original public record in which this sequence was found.

3.3.3 Additional considerations

It is possible in the main BLAST form to select other databases than the default "non-redundant" set, and select in which public database(s) to perform this search.

The "Advanced Parameters" field can be used to specify standard BLAST parameters other than the defaults used by the program. For these, help can be found by clicking the black "?" button next to the field.

4 Conclusions

This document presented a set of step-by-step procedures illustrating the different ways to use the online bioinformatics tools found on the website of the EURL GMFF. Obviously, users are invited to experiment for themselves and develop their own preferred procedures and case uses.

Request for additional clarifications, suggestions and bug reporting can be addressed via e-mail to the functional mailbox of the JRC Bioinformatics group, reachable at JRC-BIOINFORMATICS@ec.europa.eu.

References

Angers-Loustau, A., Petrillo, M., Bonfini, L., Gatto, F., Rosa, S., Patak, A., & Kreysa, J. (2014). JRC GMO-Matrix: a web application to support Genetically Modified Organisms detection strategies. BMC bioinformatics, 15(1), 1.

Bonfini, L., Van den Bulcke, M. H., Mazzara, M., Ben, E., & Patak, A. (2012). GMOMETHODS: The European Union database of reference methods for GMO analysis. Journal of AOAC International, 95(6), 1713-1719.

Patak A: CCSIS specialist EMBnet node: AGM2011 report. EMB Net J. 2011, 17 (A): 13-10.14806/ej.17.A.226.

Petrillo, M., Angers-Loustau, A., Henriksson, P., Bonfini, L., Patak, A., & Kreysa, J. (2015). JRC GMO-Amplicons: a collection of nucleic acid sequences related to genetically modified organisms. Database, 2015, bav101.

JRC101853

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